Teorías sobre la Etica

a cargo de

PHILIPPA FOOT

FONDO DE CULTURA ECONOMICA
MEXICO - MADRID - BUENOS AIRES

Primera edición en inglés, 1967
Primera edición en español, 1974

Traducción al español de
MANUEL ARBOLÍ

Cubierta: *Ruiz Angeles/Salto*

Título de esta obra en inglés:
Theories of Ethics
© 1967 Oxford University Press, Londres

D. R. © FONDO DE CULTURA ECONÓMICA
Avda. de la Universidad, 975. México, 12. D. F.

ISBN: 84-375-0008-7 (rústica)
ISBN: 84-375-0009-5 (tela)
Depósito legal: M. 36.541 - 1974

Impreso en España
Closas-Orcoyen, S. L. Martínez Paje, 5. Madrid-29

La mayoría de estos artículos fueron reimpresos sin revisión o sólo con escasas alteraciones. Así, pues, no han de reflejar necesariamente las opiniones actuales de sus autores.

I

INTRODUCCION

Los artículos que aquí se reeditan giran en torno a dos cuestiones últimamente objeto de mucha discusión: primero, la naturaleza del juicio moral y, en segundo lugar, la parte que la utilidad social tiene en determinar lo bueno y lo malo. Ambos debates se retroceden al siglo XVIII, pues en aquella sazón los filósofos andaban divididos en pro y en contra del sentido moral y de las teorías intelectualistas acerca del juicio moral. Fue también a finales de ese siglo cuando Bentham declaró que el fundamento del bien moral estaba en el principio de la utilidad.

Los demás artículos de este volumen (números del IX al XII) versan sin más sobre el utilitarismo; su referencia al pasado es, pues, clara. Los números del I al VIII no están tan abiertamente relacionados con las controversias del siglo XVIII; no obstante, su conexión es cercana. Al igual que nosotros, Hume y sus contemporáneos se sentían acuciados por la posible o imposible objetividad de los juicios morales. ¿En qué —se preguntaban— estribaba la virtuosidad de las acciones virtuosas? ¿Cómo se captaba ésta? ¿Era mediante juicio, o más bien porque se sentía? ¿Sabíamos por el entendimiento lo que se debía hacer, o por un sentido moral? ¿Había en esto algo que pudiera ser conocido, o todo discurso moral no hacía sino expresar nuestros sentires, en vez de hablar de lo que habíamos descubierto sobre la virtud o el vicio? Por su parte, Hume se convenció de que era vana la búsqueda de propiedades morales objetivas, sosteniendo

que cuando a una acción la llamábamos virtuosa no se hacía otra cosa sino sentir un sentimiento placentero de aprobación al contemplarla; teoría que parecía explicar cómo los juicios morales se vinculaban con la acción, pues naturalmente nos sentiremos inclinados a hacer o a fomentar lo que nos afecta de manera placentera, mientras que si la moralidad de las acciones residiera en algo que nos dictara la razón, se debería demostrar por qué tal descubrimiento influiría necesariamente en la voluntad.

Cabría decir que los problemas que nos inquietan hoy son precisamente los que preocupaban a Hume. Sin embargo, de manera más directa, ha sido el profesor G. E. Moore quien dispuso el tinglado en nuestro favor, no obstante que el nombre de Hume no aparece siquiera en el índice de sus *Principia Ethica*. Es como si hubiéramos empezado con Moore y hubiésemos ido retrocediendo desde él hasta Hume. Permítasenos decir algo, antes que nada, sobre las argumentaciones de vasto influjo propuestas por Moore en 1903[1]. La tesis central de Moore era que la bondad es simplemente una propiedad no-natural descubierta por la intuición. El resto de su ética se construyó sobre esta cimentación, pues Moore creía que los demás juicios morales, por ejemplo, los concernientes a la acción debida, hacían referencia a las intuiciones básicas de la bondad, resultando que la acción debida era aquella que producía la mayor cantidad posible de bien. Esta última convicción convirtió a Moore en una especie de utilitarista. Pero no ha sido esta parte de su teoría la que más ha interesado. Lo que pareció particularmente importante, al menos en las generaciones subsiguientes, fue su idea sobre el juicio que ponía en marcha todo este asunto. Sostenía Moore que estos juicios eran objetivos, pero declaraba que se producían por intuición. Por esto se le

1 G. E. Moore, *Principia Ethica*. (*V.* la bibliografía para las publicaciones que no se detallan en las notas al calce de esta «Introducción».)

llama intuicionista, compartiendo el título con filósofos como Prichard y Ross, quienes aseveraban que la intuición moral era la base del juicio moral, aunque discrepaban sobre dónde entraban las intuiciones. Es intuicionista quien cree que, al final de cuentas, no podemos dejar de 'ver' que ciertas cosas son buenas, o correctas u obligatorias. Hasta cierto punto, dicen los intuicionistas, se puede debatir una cuestión de moral demostrando que algunos casos caen bajo determinados principios por la naturaleza misma de los hechos, pero al cabo se llega a un punto en el cual no se puede decir más que 'veo que así tiene que ser'.

Las dificultades en que labora esta posición son ahora suficientemente claras, y se necesitarían muchas agallas para afirmar, cual hiciera Ross por la mitad de los años treinta, que 'todo sistema ético admite la intuición en algún punto': pues la intuición moral, a diferencia de la ordinaria, que nos advierte qué piensa o siente otro, se presume que es la 'aprehensión' de una cualidad cuya existencia no se puede descubrir por ningún otro medio. Ahora bien, si alguien sabe intuitivamente que, pongamos por caso, un individuo está enojado aunque no dé muestras de ello, dice 'me doy cuenta'; pero se sabe que alguien está enfadado por otras veces y, en principio, se puede poner a prueba las propias intuiciones buscando un medio que no deje lugar a dudas. Es así como se descubre si uno se puede fiar de las intuiciones propias, o en qué casos; a la vez que las intuiciones de algunos se pueden demostrar mejor que las de otros, porque están más estrechamente correlacionadas con los hechos objetivos independientes. Esta comprobación independiente es la que falta en las presuntas intuiciones morales, y querer reducirlas a una, hablando por ejemplo de las intuiciones que 'resisten la prueba del tiempo' o de las que tienen los 'pueblos más altamente desarrollados' es sencillamente un engaño. Pues, ¿quién nos dice que las intuiciones correctas no son aquéllas que primero pensamos

y luego abandonamos ('lo que primero viene a la mente es lo mejor')? ¿Quién nos asegura que los pueblos primitivos no poseen una facultad de intuición moral que la civilización propende a destruir?

Parece que no está justificado el recurso de los intuicionistas a la 'aprehensión' y al 'ver', dadas sus propias presuposiciones, y lo mismo vale de su afirmación de que quien 'juzga' sobre la base de su 'intuición moral' expresa una opinión acerca de algo objetivo. Puesto que si no poseemos un método que pueda decidir, siquiera en principio, entre intuiciones conflictivas, parece que no salimos de las 'trampas del corregir'. Puedo decir 'yo no tengo razón y tú estás equivocado' y 'estaba equivocado cuando dije...', pero estas proposiciones expresarán solamente una reacción, y si sólo expresan una reacción no estamos lejos de las teorías subjetivistas que rechazaban Moore y otros intuicionistas.

¿Por qué, pues, dadas las dificultades, sostuvo Moore la teoría de la intuición moral contra aquéllos que, como Hume, veían los juicios morales como expresión de los sentimientos y actitudes del interesado? Las argumentaciones de Moore en contra de esas teorías son el tema que él y el profesor C. L. Stevenson debaten en el segundo y primer ensayos de este volumen.

Defendía Moore[2] que quien afirma que cierta acción es o está correcta o equivocada no se *refiere* simplemente a que posee un sentimiento de aprobación —o cualquier otro sentimiento o actitud— hacia ella. Puesto que, según dice, ello supondría que cuando uno dijera 'X es correcto'* y otro afirmara 'X está equivocado', X estaría a la vez correcto y equivocado; y cuando alguien aseverara una vez 'X está correcto', y otra 'X está equivocado', esta misma acción individual X una vez sería correcta y otra equivocada. Objeta Stevenson que 'X está correcto' significa 'Ahora apruebo X', lo que si se aplicara *consistentemente*

[2] *Ethics*, cap. iii.

no poseería ninguna de las consecuencias de que habla Moore. Así, no podemos decir con Moore que 'Si «X está correcto» afirmado por A es verdadero, entonces X *está correcto*', y que si '«X está equivocado» afirmado por B es verdadero, X *está equivocado*'; para efecto de conclusiones, y una vez traducido, se convierte en 'Yo apruebo (desapruebo) X', pudiendo yo ser una persona diversa de A o B. No obstante, Moore posee un tercer argumento que Stevenson está dispuesto a admitir de alguna manera. Dice que la 'teoría de la actitud' subjetiva no da explicación de la discrepancia que, sin duda, se da entre dos interlocutores que, respectivamente, dicen 'X está correcto' y 'X está equivocado'. Pues si cada uno está hablando de sus propios sentimientos, ¿cómo se pueden contradecir? Uno puede tener tal sentimiento y el otro no. La respuesta de Stevenson es que, en efecto, no hay incompatibilidad lógica alguna entre las dos proposiciones: las *creencias* de los interlocutores no tienen que ser necesariamente contradictorias. No obstante, hay desavenencia entre los dos, puesto que sus actitudes son opuestas. Mas es la expresión de las actitudes opuestas la que da la oposición entre el 'X está correcto' de A y el 'X está equivocado' de B, y es sólo de esta manera como 'discrepan'.

Stevenson lucubra aquí sobre la teoría del significado emotivo de los términos éticos, asunto que se retrotrae a las discusiones del Círculo de Viena, entre 1918-19, y que claramente quedó esclarecido por Ogden y Richards cuando, en 1923, escribieron en *The Meaning of Meaning* que en lenguaje moral «...la palabra 'bueno' funge sólo como signo emotivo que expresa nuestra actitud... y que quizá evoca actitudes similares en otras personas o las incita a acciones de una clase u otra» [3]. Tal teoría había sido avanzada ya por el profesor A. J. Ayer en *Language, Truth and Logic*, pero nunca se expuso con tanto detalle como lo hiciera Stevenson en sus artículos en *Mind* de 1937

3 P. 125.

y 38, desarrollándola ulteriormente en *Ethics and Language,* publicado en 1945. Afirma allí que el sentido emotivo de una palabra es lo que la hace apropiada para propósitos tan dinámicos como la expresión de nuestras actitudes y la alteración de las ajenas, sin que posea el propósito 'descriptivo' de comunicar creencias. El significado emotivo de una palabra es su tendencia a producir respuestas *afectivas* en el oyente y a ser empleada como resultado de estados *afectivos* en el hablante.

Frente a la conclusión de que la discrepancia ética podría ser meramente actitudinal, Moore, quien de manera característica había confesado que tal posibilidad 'simplemente no se le había ocurrido', concedió que sus argumentaciones eran inconcluyentes. Así, pues, la causa del objetivismo ético parecía seguir mal curso. Como lo expresara el propio Moore, había implicado la noción de la intuición ética, con lo que habíanse desmoronado los argumentos en su favor. Mientras, fue el mismo Moore quien atacó la otra forma de objetivismo que podía haberse quedado con el campo, pues había hecho hincapié en que no podía existir definición alguna de bondad que vinculara tal propiedad con posibles cuestiones de hecho. Según esto, por ejemplo, resultaba imposible decir que 'bueno' significara meramente productor de felicidad porque se pudiera probar que ciertas cosas eran buenas. Afirmó Moore que tales teorías cometían la 'falacia naturalista'; esta vez tuvo a los emotivistas de su parte.

Que los argumentos de Moore contra el naturalismo no son concluyentes es la tesis del tercer artículo de este volumen, que se ocupa en gran parte en exponer cuáles son dichos argumentos. Piensa Moore que nadie tiene derecho a asentar proposiciones del tipo 'el placer y sólo el placer es bueno', y para ello se basa en la definición de que tales proposiciones son siempre sintéticas y nunca analíticas. Pero ¿cuál es exactamente la presunta 'falacia'? El profesor Frankena examina tres posibles opiniones: (*i*) que el

error está en definir una propiedad no natural, como la bondad, como si se tratara de algo natural, (*ii*) que la equivocación está en definir una propiedad con los términos de otra, y (*iii*) que se intenta definir lo indefinible. Arguye Frankena que sea cual sea la versión que tomemos, resulta que Moore no ha sabido mostrar que existiera error alguno y que, por tanto, no ha hecho más que una *petitio quaestionis*. Para poder asentar (*i*) debería haber mostrado que la bondad *es* propiedad no natural, cosa que solamente afirma. Respecto de (*ii*) debería haber demostrado, en cada ejemplo, que la bondad era 'algo distinto' de la propiedad con la que se equiparaba; cosa que tampoco hace. Para determinar (*iii*) debería probar que la bondad es propiedad simple y por ende indefinible, pero sólo lo asevera, sin aducir prueba alguna.

Afirma Frankena, y sin duda tiene razón en su afirmación, que Moore está convencido de que se cometía falacia naturalista con cualquier definición de bueno; pero los escritores posteriores no paran mientes en esto cuando hablan de Moore como el gran opositor de la ética naturalista. Se ciñen a excluir cierto *tipo* de definición y se remiten a lo que Moore dijo sobre la imposibilidad de identificar las propiedades naturales con las no naturales. Desgraciadamente, Moore jamás logró explicar qué entendía por propiedad 'natural'; lo más que dijo fue que la bondad de una cosa no pertenecía a su descripción, como pertenecían sus propiedades naturales. Consiguientemente, no se veía con claridad qué tipo de definición era la que debía excluirse. Sin embargo, Stevenson alegaba que su teoría del significado emotivo mostraba la verdad que Moore había buscado a tientas. El quid estaba en que la bondad no se había de tratar como una clase especial de propiedad, puesto que no era propiedad alguna; antes bien, que existía cierto tipo de significado propio de los términos éticos, y que las definiciones que omitían este elemento emotivo en el significado de 'bueno' eran de-

fectuosas. Así, pues, era posible defender el no-naturalismo de Moore, mientras que su intuicionismo estaba socavado. Se advierte que emotivistas e intuicionistas tienen algo en común: unos y otros niegan que las proposiciones morales estén abiertas a las clases de pruebas ordinarias. El intuicionista confirma que, al cabo, uno se tiene que conformar con decir 'veo que así es', mientras que el emotivista admite que será retrotraído a la expresión de sus actitudes fundamentales. Se dará fin a la argumentación una vez expuestos todos los hechos.

Durante cierto número de años fueron el emotivismo y las teorías a él concernientes el centro de atención. De estas teorías la más influyente resultó la del profesor Hare, conocida con la etiqueta de 'prescriptivismo'. Hare sustituyó el 'significado emotivo' de Stevenson, por su 'significado valuatorio' (evaluative meaning). Explicaba que cuando se empleaban 'con fuerza recomendatoria' vocablos como 'bueno' y 'debe' eran 'valuatorios' (para hacer 'juicios de valor'). Cuando se aplicaban así, comportaban imperativos, pues Hare sostiene que, por definición, si alguien emplea el juicio 'Yo debo hacer X' como juicio de valor, se ha de aceptar que '...si asiente al juicio debe tambien asentir al mandato «debo hacer X»'[4]. Así, quien emplee la palabra 'bueno' valuatoriamente, tiene que aceptar un imperativo de primera persona. Pero tras cada imperativo particular yacerá un 'cuasi-imperativo' general dirigido, por así decir, a todas las personas de todos los tiempos. Hare no está afirmando que palabras como 'bueno' y 'debe' no pueden usarse más que 'valuatoriamente', sino que su definición, de una manera u otra, parece referirse a lo que entendemos por juicio de valor en la vida de cada día. Al significado valuatorio contrapone Hare el descriptivo, pero —como Stevenson— no da razón alguna de este aspecto de su dicotomía. Para que una palabra pueda ser descriptiva no ha de ser valuatoria, y afirma que

4 R. M. Hare, *The Language of Morals*, p. 168.

deben existir 'criterios bien definidos respecto de su aplicación, en los que no se haga juicio de valor'. Una palabra puede poseer significado descriptivo y valuatorio, pero recibirá el nombre de 'palabra descriptiva' sólo si no contiene ningún elemento valuatorio.

Así pertrechado, Hare procede a lanzar un ataque por todos los flancos contra el naturalismo ético, definiendo como naturalista al que quiere equiparar las palabras valorativas con aquellas cuyo significado es 'puramente descriptivo' y que, por tanto, pretende deducir una conclusión ética de premisas descriptivas. El precio que paga el naturalismo, dice Hare, es la pérdida de la fuerza recomendatoria y de 'guía de la acción' de los términos éticos. Y propugna que una de las grandes ventajas de su propia teoría es que muestra cómo el juicio moral está conectado necesariamente con la elección. En efecto, tanto Stevenson como Hare diríase que han suministrado la conexión necesaria entre la moralidad y la voluntad, en la que había insistido Hume. En Stevenson, la conexión entre el juicio moral y la acción quedó enmarcada en la teoría del significado emotivo: el vocablo emotivo expresa las actitudes del hablante, que el oyente, en ese momento, es invitado a compartir, y puesto que toda actitud está 'marcada por estímulos y respuestas que se refieren a estorbar o a favorecer lo que se llama el «objeto» de la actitud', significa esto que el empleo de un término emotivo tiende a expresar la disposición del hablante a hacer ciertas cosas y a influir en el oyente en una dirección similar. Como hemos visto, Hare enlazó el empleo valuatorio del lenguaje con la aceptación de los imperativos en primera persona y de los cuasi-imperativos orientados al mundo en general. Por consiguiente, podía alegar que, según su teoría, los juicios de valor eran esencialmente 'guías de la acción' (action guiding), comportando esta instancia tanto respecto de las propias acciones del hablante, como de las ajenas. Partiendo de la aserción de Hume según la cual los juicios morales son prácticos necesariamente, pasó a unirla con

el famoso dictado de ese autor acerca de la brecha entre 'es' y 'debe'. No es posible deducir 'debe' alguno de las proposiciones descriptivas, puesto que los 'debe' tienen esa conexión especial con la dirección de las elecciones; lo que no ocurre con las proposiciones 'es'.

Esta posición es la que Hare defiende contra el profesor Geach en el quinto artículo aquí incluido. Geach, en su ataque, había impugnado la explicación de Hare sobre la función 'guía de la acción' de la palabra 'bueno' y su teoría del significado valuatorio. Geach acepta con Hare que 'bueno' es palabra 'guía de la acción', pues pertenece a la idea de bondad el que normalmente, y siendo iguales las demás cosas, la gente escoja aquello que recibe el nombre de bueno. Pero esto no quiere decir que, cuando se emplea en su sentido normal, dicha palabra tenga que aplicarse 'para recomendar'. En alguna ocasión particular puede darse que no se cuestione la dirección de las elecciones, en el cual caso tal palabra no se utilizará de manera especial. Así, pues, nada impide que una expresión del tipo el 'buen F' posea sentido directamente descriptivo.

A pesar de todo, Geach advierte una dificultad en su propia posición. Supongamos que la expresión 'una buena acción' posee un significado descriptivo fijo y que nos es lícito pasar —digamos— del hecho de que una acción es un acto de adulterio, al hecho de que es un acto humano malo. ¿Cómo llegaremos de la proposición presuntamente descriptiva 'el adulterio es un acto humano malo', al imperativo 'no cometerás adulterio'? ¿Por qué el pensamiento de que se trata de una acción mala habría de disuadir a alguien de cometerla? Replica Geach que 'si bien el llamar a una cosa «un buen A» o «un mal A» no es de por sí algo que toque los deseos del agente, puede ser que sí lo haga si el oyente tiene que escoger algún A'[5]. Y lo que el hombre no puede dejar de escoger es su ma-

5 *V.* p. 102.

nera de actuar; por lo que llamar a una manera de actuar buena o mala no puede sino servir para guiar la acción.

No ha de sorprender que Hare halle esta respuesta del todo insatisfactoria. Replica que si 'hombre' y 'acción' se toman como palabras funcionales, al igual que 'cuchillo' y 'soldado', entonces naturalmente 'buen cuchillo' y 'buen soldado' tendrán un 'significado descriptivo fijo'. Pero en tal caso ya no será cierto que uno no pueda dejar de escoger su manera de actuar, pues podría ser muy bien que un individuo no tuviera interés en efectuar aquellas cosas que hacen de un hombre un buen hombre, si pudiera escoger acciones que cayesen bajo otros encabezados o principios de elección. Por consiguiente, Geach no ha tomado en cuenta que el juicio moral, a diferencia de otros de la forma 'buen A', tiene que ser 'guía de la acción' para *cada* hombre, sean cuales sean sus deseos particulares. El propio Hare había garantizado esto al recalcar que 'bueno', cuando se emplea evaluativamente, conlleva en su significado una instancia a la elección; ante una palabra funcional, como 'soldado', no se emplea así, o más bien su contenido valuatorio queda neutralizado por la palabra 'soldado'. Pues esta palabra deja margen a un punto de vista especial a partir del cual es posible efectuar una elección, lo que equivale a decir 'esto debo hacer *si* quiero ser buen soldado'. Es expresión evaluativa como un todo aquélla que conlleva una regla de acción real y no hipotética, y esto es lo que sin duda debe de ser el juicio moral.

El problema que preocupa a Geach fue el que inscribía el artículo 'Creencias morales', que aparece quieto a la compiladora de la presente edición cuando con el número VI en el presente volumen. En la primera parte de dicho artículo había impugnado la idea de que en el significado de la palabra 'bueno' existiera un elemento valuatorio que fuera independiente de su significado descriptivo, alegando que no es posible extraer sentido alguno de la noción de que un hombre

piense 'esta acción es buena' si presenta pruebas cerradas para demostrar que *es* una buena acción. Ni ayuda en nada apelar a sentimientos que pudiera tener, pues hay sentimientos que no se pueden atribuir a nadie, a menos que tenga los pensamientos debidos. Esa parte de mi artículo indica que la expresión 'una buena acción' posee significado descriptivo fijo, o al menos que estaba fijado dentro de cierto margen.

Ahora bien, aunque esto haya sido rechazado por los emotivistas y prescripcionistas, que opinan que es contingente el que *nuestros* términos valuatorios posean un significado descriptivo fijo, no se trata de algo que esté exactamente en el medio de la disputa entre las dos facciones. Pues los anti-naturalistas podrían conceder que una expresión como 'buena acción' poseyera un significado descriptivo fijo, sin dejar por ello de requerir algún 'elemento volitivo' extra cuando se trate de juicios de valor. Quizá quien calificara cierta acción como buena ¿debería aplicar a ella determinadas descripciones, pero *también* poseer ciertos sentimientos o actitudes, o aceptar reglas particulares de conducta? ¿De cuál otra manera, si no, se podría mantener la fuerza 'guía de la acción'? En la segunda parte del artículo indico que esto puede muy bien ocurrir, según sean los hechos particulares con los que se relaciona la bondad de una acción buena, puesto que existen ciertos hechos concernientes a algo que dan a *cualquiera* razón para escogerlo. Tuve dificultad, desde luego, en demostrar que las acciones que consideramos como buenas son precisamente acciones de este tipo. Se puede mostrar, sin duda, que es probable que todos necesiten las virtudes del valor, de la templanza y de la prudencia, sean cuales sean sus propósitos y deseos particulares. Pero, ¿qué decir de la justicia? El ser justo no deriva obviamente en beneficio de uno mismo y puede ser que no encaje en las inclinaciones y planes de la persona.

Me hallé en esta dificultad porque presumí —con mis opositores— que el pensamiento sobre la bondad

de una acción estaba relacionado de manera asaz especial con las opciones de cada individuo. No se me había ocurrido cuestionar el dictado, frecuentemente repetido, de que los juicios morales brindan razones para la actuación de todos y de cada uno. Esto ahora me parece un error. Muy generalmente, la razón de por qué quien escoge A puede 'esperarse' que elija una buena A y no una mala A está en que nuestros criterios de bondad respecto de cualquier clase de cosa se relacionan con ciertos intereses que cada uno tiene en cada cosa. Cuando alguien comparta esos intereses tendrá razón en escoger la buena A; de otra manera, no la tendrá. Puesto que, en el caso de las acciones, distinguimos éstas entre buenas y malas, según el interés que poseamos en el bien común, a quien le importe un ardite lo que les ocurra a los demás, mientras no sea con él, podrá decir con razón que no tiene motivo alguno para ser justo. Los demás, si continuamos siendo como somos, intentaremos hacer entrar en juicio a ese individuo, diciéndole 'debes ser justo'. Es muy cierto, pues, que existen imperativos categóricos en lo moral. También es muy cierto que el 'debe' moral tiene especial fuerza 'guía de la acción', pues no se puede decir que una palabra de otra lengua es vocablo moral, a menos que pueda emplearse para *urgir* a comportarse de determinada manera. Pero esto no quiere decir que cuando se emplee para hacer otras cosas tendrá sentido diferente. Tras decir 'debes hacer X', cabe añadir sin inconveniente 'pero Dios quiera que no lo hagas'; así como también se puede decir 'debo hacerlo, ¿qué otro remedio me queda?', sin emplear la palabra 'debe' en algún sentido especial que exija las 'comillas', porque se quiera significar 'debo hacerlo', y no 'esto es lo que los demás créis que yo debo hacer'. Desde luego, tales expresiones serían excepción, pues si la gente en general no se interesara en el bien de los demás y en que se cumplieran las reglas de justicia que rigen en su sociedad, no existiría el uso moral del 'debe'. Pero de aquí no se ha de pasar a inventar un sentido es-

pecial del 'debe'. Vale decir, por tanto, que existen *dos* sentidos especiales, uno correspondiente a quien en general tomara en cuenta las consideraciones morales, pero que de vez en cuando se las saltara a la torera, y otro que se referiría a la persona amoral que jamás se fijara en lo que debe hacer.

Parece claro que todo el que rechace la idea de Hare de que los vocablos empleados para hacer una valuación han de conllevar imperativos, desechará sus argumentos particulares contra la posibilidad de deducir el 'debe' del 'es'. Soy de la opinión, por lo demás, que aquí está la verdadera cuestión candente que ventilan tanto él como el profesor Searle en los números VII y VIII de este libro. Sostiene Searle que hay al menos un ejemplo en el que cabe deducir un 'debe' de un 'es'; pues —nos dice— de ciertas premisas que nos aseveran (1) que determinadas declaraciones, hechas en circunstancias particulares, cuentan como promesas, (2) que las promesas sitúan al prometiente bajo obligaciones y que (3) Ticio profirió esas palabras en tales circunstancias, podemos sacar la conclusión —por deducción— de que, *caeteris paribus*, Ticio debe cumplir su promesa. La cláusula del *caeteris paribus* que aparece en la conclusión sirve para cautelar que las promesas no sitúan al prometiente bajo obligación absoluta, puesto que tal obligación puede quedar contrarrestada por otras consideraciones, cual una obligación procedente. Pero, asimismo, esa misma cláusula puede inserirse en las premisas, con lo que resulta nueva premisa que asevera que hay igualdad de condiciones, deduciéndose una conclusión simple (no condicional) sobre lo que Ticio debe hacer. Muchos de los debates de ese artículo están centrados en el *caeteris paribus*, no así en el de Hare; me parece que tiene razón en pensar que no es el punto clave. Si a Hare se le presentara el caso de que, mediante hechos, se hubiera extraído un 'debe' de un 'es' respecto de una instancia como la promisión, replicaría de la manera siguiente. Diría: o bien 'tengo obligación de guardar mi promesa' es una

proposición prescriptiva, o no lo es; es decir, o conlleva un imperativo de primera persona o no lo conlleva. Si no es prescriptiva, no es valuatoria y, por tanto, no se ha podido deducir una conclusión valuatoria de premisas que son puramente fácticas. Por el contrario, el que sea prescriptiva no se puede deducir de proposiciones descriptivas de este tipo, pues la cuestión es si yo, el hablante, me someto a las reglas del 'juego' del prometer. Sin duda, la existencia de la institución del prometer requiere que haya algunos que acepten esas reglas, pero tal hecho 'antropológico' no liga mi voluntad, y de él sólo podría deducir otro hecho 'antropológico'[6].

Searle respondería, sin duda, que el 'debe' por él deducido no es valuatorio en el sentido de Hare, pues niega que las proposiciones descriptivas y las valuatorias se puedan distinguir, como supone éste. Propone que, en vez de buscar algún tipo especial de significado en las 'declaraciones valuatorias', deberíamos atender ante todo a las múltiples cosas (evaluándolas recíprocamente) que podemos hacer *al* usar una forma particular de palabras. Searle emplea aquí la distinción que hace el profesor J. L. Austin entre la 'fuerza locucional' de una expresión, que más o menos equivale a su significado, y el 'acto ilocucional' que el hablante puede llevar a cabo al decir lo que hace[7]. La evalución sería sin más uno de los muchos actos ilocucionales que se puede hacer ejecute una forma dada de palabras.

Presumiblemente, Searle echaría mano de esta misma distinción entre significado y acto de proferir para responder a la objeción central de Hare a su argumentación. Según Hare, la cuestión crucial está en si podemos o no podemos, con Searle, considerar como tautología que 'Bajo ciertas condiciones C, todo el que profiera las palabras (proposición) «Con esto

6 *V.* p. 179.

7 Prometer, advertir, suplicar, recomendar, conminar, reconvenir,

te prometo, Cayo, pagarte cinco pesos» se coloca bajo (asume) la obligación de pagar a Cayo cinco pesos'. (Se trata de saber, para decirlo brevemente, si es una tautología el que las promesas se deben cumplir, como había dicho Searle que así era.) Afirma Hare que si fuera una tautología, no podría estatuir una regla del 'juego' del prometer, puesto que si la estatuyera debería decir *cómo actuar*. En otras palabras, quiere indicar Hare que en las palabras que establezcan una regla debe existir un elemento prescriptivo. A lo que podría replicar Searle que si bien la palabra 'debe', según se emplea en determinadas circunstancias, posee en verdad la fuerza ilocucional de *mandar*, no quiere decir que haya una *consecución* extra que permita pensar en un argumento deductivo de 'es' a 'debe'.

No sé decir si me he apartado de Searle al inventarle esta réplica, ni si es así como se puede explicar esto. En todo caso concuerdo con Hare en hallar defectuoso el argumento de Searle, aunque mis razones son harto diferentes de las suyas, pues me parece que si bien, *en principio*, nada obsta que se intente derivar 'debe' de 'es', Searle ha operado con premisas de mala calidad, al menos por lo que hace al 'debe' moral. Ha querido deducir una proposición 'debe' de premisas que son 'internas' de una institución particular, y no es así como se emplean las proposiciones 'debe'. Para ver esto no tenemos más que suponer que poseemos una institución del todo mala —digamos, una institución que se refiera al duelo— a tenor de cuyas reglas uno tiene obligación de disparar a otro, una vez que se han dicho y hecho ciertas cosas. En tal caso podríamos construir un argumento paralelo al de Searle que conduciría a la conclusión de que hay que disparar contra X. Pero, de hecho, nadie que reprobara tal institución fundándose en razones morales afirmaría eso. Negaría que tuviera obligación alguna de disparar contra su contrincante, debido a las aviesas consecuencias que esa institución

tiene para la sociedad; no precisamente porque repudia obedecer la reglamentación (de lo que se puede o no se puede tratar), sino porque denegara que existiera la *obligación*, por su manera de ver la institución. Vale decir que mientras Searle no andaba errado en principio al afirmar que se podía derivar el 'debe' del 'es', descuidó pensar que se pudiera inferir de esas premisas particulares. Puesto que, si bien algunas palabras que naturalmente pueden recibir la denominación de 'evaluativas' (*e.g.* 'adeudar') parecen pertenecer a una institución [8], 'debe' sólo se puede deducir de un conjunto de premisas que hacen referencia a cosas como la ofensa, la libertad y la felicidad; es decir, instancias que cuentan en la escala del bien o del daño humanos. Así, no es posible negar en verdad que uno adeude cierta suma de dinero, de acuerdo con determinadas instituciones y asuntos institucionales de hecho, según el tipo que Searle tiene en mente; pero si alguien considerara que toda la institución era perjudicial y que destruirla fuera cometido social provechoso, diría 'no es verdad que se deba pagar lo que se adeuda'. Según esto, 'hay que cumplir las promesas' no es una proposición tautológica, y todo lo más que se puede decir es que la promesa presupone la aceptación de una obligación por parte de cierto número de personas. Por lo que respecta a la deducción de 'debe' de 'es', se habrá de ver que las premisas sean correctas y esperar qué resulta. Hare no ha demostrado que en principio se pueda objetar a tal procedimiento, ni Searle ha probado que se pueda hacer. Todo dependerá de cómo se relaciona el significado de 'debe' en un juicio moral, con nociones referentes al perjuicio o al bienestar, y esto se ha de elaborar todavía.

Si uno contempla los últimos veinticinco años, queda sorprendido y algo triste porque este conflicto particular sobre el 'hecho y el valor' ha ocupado tanta parte de nuestra época. Parece como si hubiésemos

8 *V.* G. E. M. Anscombe, «On Brute Facts», *Analysis* (enero, 1958).

irrumpido en el campo sin esperar a trazar el mapa del territorio en disputa, dispuestos a morir por ciertas tesis sobre la recomendación o la aprobación, sobre actitudes en pro o sobre valuación, antes de que nadie realizara alguna tarea detallada acerca de los conceptos específicos y tan diferentes que entraban en el asunto. De hecho, la filosofía moral se ha beneficiado *relativamente* poco de la revolución que en otros campos nos ha puesto a parar mientes en el lenguaje de cada día, así como de la más o menos paciente investigación del detalle. Es raro, por ejemplo, que no fuera sino hasta la tardía fecha de 1956 cuando Geach sostuviera que la valuación no se podía representar por el en general sin sentido 'X es bueno'. Y es raro que no se haya trabajado más sobre conceptos como los de la actitud y sobre las diferencias pequeñas (¿o grandes?) entre aprobar, recomendar, encarecer, advertir [9], elogiar, valorar y semejantes. Será natural que nos volquemos sobre esos temas ahora que Austin nos ha mostrado algunos modos para hacerlo. Se siente que esta parte de la filosofía moral va a tener que cambiar en bien, una vez que se haya asimilado totalmente su obra. El propio Austin afirma que 'el contraste familiar entre «lo normativo o evaluativo», en cuanto opuesto a lo fáctico, al igual que tantas otras dicotomías, necesita ser eliminado' [10], y todo podría ser que nos percatáramos de que hemos ido haciendo demasiadas contraposiciones cuando bastaba con una.

II

En los artículos impresos con los números del IX al XII de este volumen, el señor Urmson, el señor

9 Pero *V.* D. J. Diggs, «A Technical Ought», *Mind* (1960).
10 Austin, *op. cit.*, p. 148.

Mabbott, el profesor Smart y el profesor Rawls dilucidan cierto problema referente a la interpretación y defensa del utilitarismo en ética. Rozan, por tanto, la tesis sobre que las acciones se pueden convertir en buenas o malas según sean buenas o malas las consecuencias; pues podemos aceptar tal cosa como definición general del utilitarismo, dejando abierta la cuestión de si se han de identificar las consecuencias buenas con la mayor felicidad del mayor número, como querían Bentham y Mill, o si, con Moore, debemos suponer que hay otros bienes últimos, además de la felicidad. No se discuten aquí esas distinciones, sino que se tratan ciertas dificultades a que se enfrentan los dos tipos de utilitaristas que quieren reconciliar el principio general que juzga las acciones por su utilidad, con los juicios morales que de hecho practica la gente. Algunos de éstos son particularmente arduos; así, por ejemplo, pensamos normalmente que existe cierta obligación de cumplir las promesas, lo que no depende de la utilidad que se recabe. Pues si bien alguien puede, alguna vez, quedar absuelto de cumplir una promesa por el daño que resultaría de satisfacerla, no nos sentimos inclinados a considerarnos absueltos nosotros por el mero hecho de que el cumplimiento de la promesa no traería ningún bien o porque romperla reportaría más bien que perjuicio. Más aún, es razonable sostener que existen ciertas acciones que ninguna consecuencia buena justificaría, *v. g.*, la tortura o la condenación judicial del inocente; y aun aquéllos que conceden que, en algunas contingencias, incluso esas cosas pudieran justificarse, de ordinario desechan la idea de que tuviéramos derecho a fingir secretamente un juicio y ahorcar a un hombre inocente para salvar por ese medio la vida de otros dos. Más aún, que no consideramos lícito sacrificar a los deficientes mentales en aras de la investigación médica.

Para solventar estas dificultades se ha propuesto aplicar la prueba de utilidad no a las acciones indi-

viduales, sino a tipos de acción, según lo cual no deberíamos preguntar '¿tendrá buenas consecuencias romper esta promesa (confabularse contra el inocente)?', sino más bien '¿resultaría bien o mal de la regla que permitiera no cumplir las promesas (o incriminar al inocente)?'. Si las consecuencias de tal regla fueran malas, también lo sería la acción individual que cayera dentro de esa regla, aun cuando *sus* secuelas fueran buenas.

Es una versión de esta teoría, a veces llamada 'utilitarismo regular' en oposición al 'utilitarismo de los actos', y a veces 'utilitarismo restringido' contraponiéndolo al 'utilitarismo extremo', la que Urmson atribuye a Mill. No afirma que fuera un utilitarismo regular a carta cabal, puesto que Mill dice que en ciertos casos se ha de aplicar la prueba de las consecuencias directamente a las acciones individuales, a saber, cuando hay conflicto entre las reglas o cuando no se puede aplicar regla alguna; sino que Urmson opina que Mill respondería a algunas objecciones al principio de la utilidad recalcando que es la tendencia de un tipo de acción lo que cuenta. Mabbott cuestiona esta interpretación de Mill, y lanza objeciones respecto de la racionalidad de tal regla. Smart va más adelante, aseverando que sería irracional hacer algo que pugnara con el principio de la utilidad aplicado a los casos individuales. ¿Por qué nos tendríamos que preocupar por los resultados que nuestra acción pudiera tener en otra parte, si no los tiene aquí? Smart se declara utilitarista extremo, opinando que si nuestros juicios morales se oponen al principio de utilidad, tanto peor para ellos.

Por otra parte, Rawls piensa que la aplicación 'utilitarista regular' del principio de utilidad es lícita en ciertos casos y cree que ello ayuda a resolver las dificultades en que incurre el utilitarista. De los cuatro artículos, éste es el más complejo y merece comentario especial. Antes que nada hay que aclarar que no se puede llamar a Rawls 'utilitarista regular' a menos que se especifique bien, pues en artículo pos-

terior [11] ha hecho constar que no cree que exista versión alguna del utilitarismo que sea compatible con todos los principios de la justicia. Por tanto, no se adhiere a ninguna de las formas del principio de utilidad, aunque cree que es lícito contender en pro y en contra de ciertas acciones, basados en motivos utilitaristas, y que en ciertos casos muy especiales se ha de aplicar el utilitarismo regular. Estos casos especiales son aquellos en los que interviene una actividad (como *v. g.* la promisión o el castigo) cuya existencia depende de reglas de acción que no permiten a la persona decidir qué ha de hacer ponderando simplemente las consecuencias. Señala Rawls que no existiría la promisión o el castigo en un mundo en el que cada cual hiciera lo que juzgara reportaría las mejores consecuencias en cada caso particular, puesto que una promesa impone ulteriores restricciones a lo que uno ha de hacer, y la punición ha de conformarse a ciertas reglas que versan sobre las injurias y penas. Así, pues, las instituciones del castigo o del prometer presuponen una conducta que en este sentido no es utilitarista.

Rawls saca la consecuencia de que la justificación de toda acción que presuponga tales prácticas (*v. g.*, la ruptura de una *promesa*) debe conformarse a las reglas de la institución, de manera que se habrán de tener presente las consecuencias sólo hasta el punto en que las reglas lo permitan [12]. Es la práctica y no la acción individual la que ha de resistir la prueba utilitarista. Lo que desconcierta es por qué Rawls piensa que sea posible extraer tal conclusión. Arguye Smart que una persona que pudiera quebrantar las reglas sin dañar la institución útil, actuaría irracionalmente si no lo hiciera cuando las consecuencias fueran buenas, y contra esto no parece que Rawls haya presentado defensa alguna. Una cosa es mostrar que las reglas que rigen cierta práctica tienen que ser

11 J. Rawls, «Justice as Fairness», *Philosophical Review* (1958).
12 *V.* p. 239 y 240.

no-utilitaristas, y otra convencer de que un individuo no puede apelar secretamente al principio de la utilidad, contrariando las reglas.

Finalmente diremos alguna palabra sobre la relación entre los problemas presentados en los dos grupos de artículos, I-VIII y IX-XII. Son de diversas clases, puesto que Moore, Stevenson, Frankena, Geach, Hare, Foot y Searle, hablan acerca del carácter lógico del juicio moral, mientras que Urmson, Mabbott, Smart y Rawls tienen en mente la interpretación y adecuación de determinado criterio referido a correcto o errado. Estos últimos no dicen nada sobre el *status* o calidad del criterio y dejan sin ventilar si se ha de considerar al utilitarista (sea utilitarista de los actos o regular) como intuicionista, emotivista o naturalista. Podría ser cualquiera de las tres cosas, puesto que, dada cierta versión del principio de utilidad —a saber, que 'las acciones son correctas mientras tiendan a producir la mayor felicidad para el mayor número'—, podría considerarse que se trata sea (a) de un juicio practicado por intuición, (b) de una expresión actitudinal o (c) de una verdad analítica. Por lo tanto, el decidirse por el utilitarismo o contra él no compromete a nadie a adoptar posición particular alguna respecto de las teorías intuicionistas, emotivistas o naturalistas de la ética y, de manera similar, los intuicionistas, los emotivistas y los naturalistas son igualmente libres de aceptar o rechazar el principio de la utilidad.

I

ARGUMENTOS DE MOORE CONTRA CIERTAS FORMAS DE NATURALISMO ETICO

C. L. Stevenson

De *The Philosophy of G. E. Moore*, a cargo de P. A. Schilpp, Volumen IV de la 'Biblioteca de Filósofos aún Vivos' (Library of Living Philosophers), Northwestern University Press, Evanston, Ill., 1942), pp. 71-90. (Se harán las futuras ediciones por Open Court, La Salle, Ill., y por Cambridge University Press, Londres). Este artículo se reimprime con permiso de Library of Living Philosophers, Inc.

En el tercer capítulo de sus *Ethics* [1], Moore presentó varios argumentos para demostrar que 'ser o estar correcto' o 'estar equivocado' no se refieren meramente a sentimientos o actitudes de quien hace uso de esos conceptos. Durante los treinta años transcurridos desde entonces, Moore se ha vuelto más sensible a la flexibilidad del lenguaje ordinario, por lo que dudo de si todavía mantendría que *nunca* se ha de emplear 'estar correcto' y 'estar equivocado'; pero quizá sostuviera que *si* alguien utiliza estas expresiones en esa forma, lo ha de hacer en un sentido que no se relacione con las instancias que de ordinario emplean los moralistas. Al interpretar algunos de sus argumentos de modo que aparezca en ellos esta segunda actitud, me propongo determinar qué es lo que prueban. *Lo que tales argumentos propugnan*, expresado de manera más formal, es que las definiciones,

1 Henry Holt & Co., N. Y., 1912.

D_1: 'X está correcto' tiene el mosmo significado que 'Yo estoy de acuerdo con X', y que
D_2: 'X está equivocado' equivale a 'Yo estoy en desacuerdo con X' [2],

—donde el 'Yo' del *definiens* se ha de tomar como refiriéndose a quienquiera que emplee los términos definidos— son definiciones que distorsionan o pasan por alto los sentidos que más importancia poseen en la ética normativa.

Si los argumentos de Moore lograran probar esta propuesta, serían de interés sin duda alguna. Tiene que haber cierta razón más o menos clara, o conjunto de razones, para que no sean sólo autores profesionales de ética normativa, sino también 'moralistas aficionados' de todo tipo quienes se empeñen con esmero en determinar qué es lo correcto o lo incorrecto, y discutan con otros estos temas. Todas esas personas recibirían buena ayuda de definiciones que liberaran de confusiones el empleo de 'correcto' y 'equivocado'. Contrariamente, ningún auxilio alcanzarán de definiciones que refieran esos términos a algo totalmente extraño a los aspectos que, por confusos que puedan ser, causan desconcierto en ellas. Si D_1 y D_2 actuarán así y si fueran insertas con persistencia en una argumentación corriente sobre ética, sólo lograrían 'cambiar el contenido' de la discusión, aunque de forma que escaparía a la atención, porque se emplearían aún las palabras de antes; serían definiciones con 'petición de principio'.

Naturalmente, hay respuesta a esta coyuntura. El teorizante puede replicar que el modo como la gente emplea 'correcto' y 'equivocado' es *del todo* confuso y no es posible poner a salvo instancia alguna en el

2 Las palabras 'estar de acuerdo' y 'estar en desacuerdo' pueden entenderse como designativas de sentimientos que el hablante *tiende* a poseer, lo que le permite hablar confiado en que dice verdad (truthfully) acerca de si tiene acuerdo o desacuerdo actual, incluso aunque en el momento no tenga sentimientos inmediatos y fuertes. Moore ha hecho referencia a esto con relación a Westermarck, en *Philosophical Studies*, 332.

tipo de argumentación ordinaria de ética. Luego, podría *conceder* a los términos un significado que estuviera de acuerdo con D_1 y D_2, sin esperar ser 'fiel' a las confusiones del uso común, aunque pretendiendo obligar a la gente a percatarse de que si no emplea su sentido, u otros sentidos naturalistas como el suyo, no hará sino tratar pseudoproblemas. De igual manera, el behaviorista definirá 'alma' como procesos que tienen lugar en el alto sistema nervioso; con ello (tómese como se quiera), pretenderá probablemente que la gente crea con él que 'alma' o ha de significar algo así o no es más que la etiqueta de algo confuso.

Se *puede* proceder de esa manera, pero no es mi intención hacerlo. Aunque los términos de ética se emplean de guisa manifiestamente confusa, no hay motivo para proclamar que existirá 'confusión *total*' a menos que se prueben cuidadosamente todas las opciones. Para empezar, no estará mal suponer que los términos de ética, como se emplean de ordinario, *no* son del todo confusos. Tal presuposición nos llevará a *mirar* si existe algún elemento salvable en su empleo. Si no miramos, no sabremos si existe ni si es precisamente ese elemento el que da a la ética normativa una de sus dificultades más características. Así, pues, presumamos siquiera, por ahora, que los términos éticos no están del todo confusos y, además, que *si* los argumentos de Moore prueban debidamente su tesis —si D_1 y D_2 distorsionan o pasan por alto los sentidos que más interesan a los autores de temas morales—, entonces se trata de definiciones que piden cuestión y que producen mayor confusión, en vez de ser enfoques dilucidadores.

El primer argumento se puede formular, sin alteración notoria de la fuerza de las propias palabras de Moore [3], como sigue:

[3] *Ethics*, 91: 'Si cuando juzgo que una acción está correcta no hago sino juzgar que yo poseo un sentimiento particular hacia ella, entonces se sigue llanamente que, con tal de que en realidad posea tal sentimiento, mi juicio es verdadero y, por tanto, la acción en

(1) Puede suceder que un hombre, A, esté de acuerdo con X, y que otro hombre, B, esté en desacuerdo con X.

(2) Así, según D_1 y D_2 —arriba—, A puede decir que 'X está correcto' y B, 'X está equivocado', y ambos decir verdad[4].

(3) Por tanto, si 'correcto' y 'equivocado' se emplean de consonancia con D_1 y D_2, X puede estar correcto y equivocado a la vez.

(4) Pero si 'correcto' y 'equivocado' se emplean en algún sentido ético típico, entonces X no puede estar a la vez correcto y equivocado. (Esto queda patente mediante la 'inspección'[5].)

(5) Consiguientemente, el sentido que D_1 y D_2 dan a 'correcto' y 'equivocado' no es sentido ético alguno.

La crítica del primer argumento tiene que realizarse de algún modo que sea de la incumbencia del razona-

cuestión realmente está correcta. Y lo que a este respecto valga para mí, valdrá también para cualquier otro... Se sigue estrictamente, por ende, de esta teoría, que siempre que *cualquier hombre* posea realmente un sentimiento particular respecto de una acción, la acción en verdad estará correcta, y siempre que *cualquier hombre* posea realmente otro sentimiento particular respecto de una acción, tal acción es en verdad errónea'. Y, 93: 'Si tomamos en cuenta un segundo hecho, parece seguirse claramente que... con harta frecuencia una misma acción puede estar correcta y equivocada. Este segundo hecho es, sin más, el hecho observado —que parece difícil denegar— por el que, sean cuales sean los pares de sentimientos o el sentimiento singular que tengamos, ocurrirían casos en que dos hombres diferentes experimentarán sentimientos opuestos respecto de la misma acción.'

[4] A tenor de los convencionalismos usuales en lógica, la 'X' no puede sufrir sustitución alguna, si aparece entre comillas. Aquí, no obstante, no tengo inconveniente en que 'X' se emplee de distinta manera. Si el lector borrara el símbolo 'X', aparezca o no aparezca entre comillas, y lo sustituyera *del todo* por el nombre de una acción particular, habida cuenta de que el nombre fuera perfectamente inequívoco, seguiría teniendo el tipo de argumento que busco. Con esta explicación se captará mejor qué quiero dar a entender al decir que 'X está correcto' puede ser verdadero. Quiero significar simplemente que esa expresión, al cambiarse su primera letra por un nombre, puede decir verdad.

[5] *Ethics*, 86 y *s*.

miento de Moore (3). ¿Es posible, mediante premisas inocentes y lógica válida, probar que si 'correcto' y 'equivocado' se emplean de acuerdo en D_1 y D_2, X puede estar a la vez correcto y equivocado? Podemos sospechar con razón que no es posible, simplemente porque de D_1 y D_2 se puede derivar una conclusión del todo diferente. Así, la última parte de (3),

(*a*) X puede estar a la vez correcto y equivocado, se convierte en equivalente, por D_1 y D_2 (como puede verse por simple sustitución con cambios gramaticales insignificantes), a

(*b*) Yo puedo estar conforme y disconforme a la vez con X. Esta última proposición puede tomarse, dentro de los límites de la propiedad lingüística, como una contradicción. Por ende, D_1 y D_2 implican que (*a*) se puede tomar como una contradicción. A tenor de esto se puede urgir que

(3x) Si 'correcto' y 'equivocado' se emplean de acuerdo con D_1 y D_2, X no puede estar a la vez correcto y equivocado. Adviértase que esta conclusión, lejos de señalar algún camino en que D_1 y D_2 distorsionen el empleo común, nos indica que una y otra son fieles a éste. Adviértase, asimismo, que si debemos admitir tanto (3x) como (3) de Moore, hemos de concluir que D_1 y D_2 implican la contradicción de que X puede y también posiblemente no puede estar a la vez correcto y equivocado. Ahora bien, ¿distorsionan D_1 y D_2 el empleo ordinario? Es difícil mantener que definiciones tan inocentes contengan contradicción tan flagrante. Por tanto, si aceptamos la derivación de (3x), tendremos motivos para sospechar con razón que existe algún error en la derivación mooreana de (3).

No es preciso, claro está, sostener que (*b*) —arriba— es una contradición. Y como habitualmente propendemos a formar sentidos consistentes con cualquier declaración, podemos llegar sin dificultad a interpretaciones más caritativas. Podemos tomarlo como un modo paradójico de decir 'Yo puedo estar de acuerdo con ciertos aspectos de X y también discordar de

otros'; o podemos considerarlo como testimonio de un posible conflicto de actitudes; como si fuera un modo paradójico de decir 'Ciertos impulsos míos me pueden conducir a aprobar X, mientras que otros me pueden guiar a lo contrario'. Pero si accedemos a hacer estas interpretaciones más caritativas de (*b*), ¿no es posible que dejemos de hacerlas con (*a*) y, por tanto, proceder a cuestionar (4)? Si existe alguna razón contra esto, Moore no la menciona. En todo caso existe ciertamente un medio, lingüísticamente apropiado, de interpretar (*b*) como contradicción; así, pues, hay por lo menos una opción en el uso del *definiens* en que D_1 y D_2 no se ha visto que distorsionen el empleo ordinario. Las definiciones pueden ser todavía objetables, pero el primer argumento de Moore en manera alguna ha demostrado que lo sean.

Es interesante ver dónde es inválida la derivación mooreana de (3), según mi versión que, a mi manera de ver, es fiel. Este paso parece seguirse de (2), que a su vez es perfectamente exacto; pero parece seguirse sólo por cierta confusión en los pronombres [6]. En (2), donde se lee 'Según D_1 y D_2, A puede decir que «X está correcto» y B, que «X está equivocado», y ambos decir verdad', las palabras 'correcto' y 'equivocado' son citas directas. Por ende, el vocablo 'Yo' que, por D_1 y D_2 va implícito en el uso de los términos éticos, se supone debidamente que no se refiere a Moore o a cualquier otro hablante, sino a la gente que se dice ha afirmado que X estaba correcto o equivocado. El 'Yo' implícito en 'correcto' se refiere a A y el 'Yo' implícito en 'equivocado' se refiere a B. Pero en (3), que puede abreviarse como 'Según D_1 y D_2, X puede ser a la vez correcto y equivocado', las palabras 'co-

6 Esta confusión a menudo requiere el empleo de lo que el doctor Nelson Goodman ha denominado 'palabras indicadoras'. En gran parte, mi crítica del primer argumento de Moore exige la aplicación del trabajo de Goodman a un caso especial. Véase el Cap. XI de su *A Study of Qualities*, disertación doctoral que ahora sólo se puede conseguir en Widener Library, Harvard, pero que será publicada por Harvard University Press.

rrecto' y 'equivocado' no las cita Moore como si las hubiera dicho algún otro. Por tanto, según D_1 y D_2 —que señalan que sus términos éticos se refieren al hablante que los emplea (para distinguirlo del hablante que cita cómo los emplearon otros)— el 'Yo' implícito en (3) no se refiere en primer lugar a A y, luego, a B, antes bien a Moore o a quienquiera que sea el que diga que 'X puede estar a la vez correcto y equivocado'. Para decirlo más brevemente, los 'Yos' citados implícitamente en (2) no se refieren a la misma persona a que se refieren los 'Yos' implícitos y no citados de (3). Al suponer que sí se refieren, Moore hace que parezca válido un paso en falso de su argumento.

Este particular se puede esclarecer exponiéndolo de otra manera. Parecería que

(a1) Si 'X está correcto', dicho por A, es verdadero, entonces X está correcto; y que

(a2) Si 'X está equivocado', dicho por B, es verdadero, entonces X está equivocado.

Y es ciertamente verdadero que *si* (a1) y (a2) son verdaderos entrambos, y *si* sus antecedentes pudieran ser entrambos verdaderos, entonces sus consecuencias serían verdaderas a la par. Así, si D_1 y D_2 permitieran aceptar (a1) y (a2) y a la vez dieran como posible la conjunción de sus antecedentes, franquearían concebir como posible la conjunción de sus consecuentes o, en otras palabras, aseverar que X podría estar a la vez correcto y equivocado. Esto es lo que, según (3), parece sostener Moore en parte. Pero desgraciadamente para la argumentación de Moore, D_1 y D_2 no legitiman aceptar *ni* (a1) *ni* (a2). Pues, por D_1, (a1) equivale a:

Si 'Yo estoy de acuerdo con X', dicho por A, es verdadero, entonces acepto X.

Y, por D_2, (a2) equivale a:

Si 'Yo estoy en desacuerdo con X', dicho por B, es verdadero, entonces repudio X.

Mas ninguna de estas proposiciones es verdadera mientras los 'Yos' citados en los antecedentes tenga

cada uno factor referente distinto del de los 'Yos' no citados en los consiguientes. Se ve así que Moore, que presume tácitamente (a1) y (a2) al pasar del paso (2) al (3) en su argumento, no logra mostrar que D_1 y D_2 llevan a lo que, según el uso ordinario, sería un absurdo. Al querer dejar patente el absurdo, rechaza —sin percatarse— que exista implicación de estas definiciones en la falsedad de (a1) y (a2), y, de esta guisa, *rechaza* las definiciones en el decurso mismo de una argumentación que trata de demostrar el absurdo que supondría su *aceptación*.

Si D_1 y D_2 se leyeran, respectivamente:

'X está correcto' tiene el mismo significado que '*Alguien* está concorde con X' y

'X está equivocado' tiene el mismo significado que 'Alguien está en desacuerdo con X', donde el 'alguien' podría ser persona diferente en cada caso, Moore podría haber pasado al (3), y su argumento hubiera demostrado que *esas* definiciones naturalistas distorsionan el uso ordinario, toda vez que se conceda (4). Pero con demostrar esto meramente, dejaría sin tocar las definiciones mucho más interesantes que nos proporcionan D_1 y D_2.

Aunque no en palabras de Moore[7], si bien en expresiones que sin duda son fieles a D_2, A puede decir 'X está correcto', y B, 'X está equivocado' y ambos decir verdad. Habría podido ser que Moore hubiese procedido por *otro camino*, a partir de este punto, para demostrar que estas definiciones violan el uso ético ordinario. Creo, sin embargo, que la única senda plausible es la que el propio Moore explana en su tercer argumento, que aquí hemos alistado y que discutiremos en su lugar apropiado.

El segundo argumento formulado, de igual manera, no en palabras de Moore[7], sino en otras que indudablemente son fieles a su contenido, dice:

7 *Ethics*, 97: 'Una acción... [que alguien] antes consideró con... repudio, puede ahora mirarla con... aceptación, y *viceversa*. Así,

(1) A puede decir verdad si afirma 'Yo ahora apruebo X, pero antes discordé de X'.

(2) Por tanto, por D_1 y D_2, A puede estar diciendo la verdad si asevera 'X ahora está correcto, pero antaño estaba equivocado'.

(3) Pero en cualquier sentido típicamente ético de 'correcto' y 'equivocado', A *no* puede decir verdad al afirmar 'X ahora está correcto, pero antaño estaba equivocado'. Podría ser cierto, a lo más, si cada 'X' se refiriera a una acción *diferente*, aunque del *mismo tipo*, que tuviera diversas consecuencias según se tratase de la X presente o de la anterior. Pero habría contradicción en cualquier sentido ordinario de los términos si 'X' se refiriera siempre, como es la intención aquí, a la misma acción. (Esto puede verse por 'inspección'.)

(4) Así, pues, el sentido adscrito a 'correcto' y 'equivocado' en D_1 y D_2 no es sentido ético típico en modo alguno.

pues, por esta sola razón, e independientemente de las diferencias de sentimientos entre las distintas personas, habremos de admitir, de acuerdo con nuestra teoría [a saber, las definiciones criticadas en el argumento en cuestión], que con frecuencia *ahora* es verdad de una acción que *estaba correcta*, aunque era primeramente verdad de la misma acción que *estaba* equivocada.'

He tratado de conservar la fuerza de estas palabras en los pasos 1) y 2) de mi formulación del argumento. Será patente que me he tomado licencias, pero las palabras de Moore se vuelven tan intrincadas, por lo que hace a los tiempos de los verbos, lo mismo que con 'primeramente' y 'ahora' y la noción de 'verdad en un tiempo pero no en otro', que sería imposible indagar más cabalmente en lo que quiere decir en espacio limitado. Goodman analiza exhaustivamente, aunque sin hacer referencia a Moore, la noción de 'verdad en un tiempo', y otras fuentes de confusión, en la nota 6 de la obra antes citada [nota 6]. El lector que se interese en estas cuestiones hará bien en acudir a dicha obra. En el ínterin, sólo puedo recalcar que si hubiese sido más fiel a las palabras de Moore, me habría encontrado ante más falacias a desenredar que las contenidas en la formulación actual de su argumentación.

Los pasos 3) y 4) de mi formulación son paralelos a las advertencias de *Ethics*, 86 y 81 *ss.*

En la crítica del segundo argumento se ha de atender a la derivación del paso (2). Este parece seguirse directamente de (1) por sustitución, de acuerdo con D_1 y D_2, pero de hecho requiere también de 'corolarios', por así decir, de D_1 y D_2; a saber:

D_{1c}: 'X estaba correcto (antaño)' tiene el mismo significado que 'Yo (antaño) concordaba con X' y

D_{2c}: 'X estaba equivocado (antaño)' tiene el mismo significado que 'Yo (antaño) discrepaba de X'.

Estas definiciones difieren de D_1 y D_2 sólo en que la referencia temporal, tanto en el *definiendum* como en el *definiens*, cambia del presente al pretérito[8]. Es obvio, sin más, que (2) se sigue de (1), si se concede que D_1 y D_2 poseen los 'corolarios' de arriba, y puesto que acepto el resto del argumento (aunque no sin dudas respecto de (3)), acepto [todo] el argumento. Pero sólo con la condición de que se entienda que D_{1c} y D_{2c} están incluidas en D_1 y D_2.

Ahora bien, es cosa del todo natural suponer que D_1 y D_2 *incluyen* D_{1c} y D_{2c}. Pero existe otra posibilidad que no deja de tener interés. Se puede insistir en que 'correcto' y 'equivocado' se refieren siempre a las actitudes que tiene el hablante *en el momento en que usa las palabras*. Cualquier referencia temporal en toda proposición que contenga esas palabras se puede tomar siempre como que hace referencia al tiempo en que *ocurrió* la acción que se dice está 'correcta' o 'equivocada', y no al tiempo en que se *aprobó*. Tal manera de ver las cosas queda explanada en las siguientes definiciones, que son versiones corregidas de D_1 y D_2:

8 De hecho, sólo D_{2c} es la que se requiere para la inferencia de 1) a 2), junto con D_1. Pero enlisto también D_{1c} simplemente porque el argumento podría refundirse muy fácilmente de una manera que la requeriría.

D_3: 'X { está / estaba / estará / estaría / etc. } correcto', tiene el mismo significado que 'Yo ahora estoy de acuerdo con X, que { está / estaba / estará / estaría / etc. } ocurriendo'.

D_4: 'X { está / estaba / estará / estaría / etc. } equivocado' tiene el mismo significado que 'Yo ahora discrepo de X, que { está / estaba / estará / estaría / etc. } ocurriendo'.

Adviértase que con estas definiciones no se puede decir nada que sea equivalente a 'Yo estuve de acuerdo con X' sirviéndose de 'correcto', a menos que, en todo caso, se emplee un giro como 'Solía sentir que X estaba correcto'.

Es fácil ver que si el segundo argumento se redactara ahora con referencia a D_1 y D_2, pero reemplazándolas con D_3 y D_4, no tendría validez. (2) no se seguiría de (1), pues la proposición X (ahora) está correcto, pero X estaba antaño equivocado sería equivalente —según D_3 y D_4—, por sustitución directa a: Yo ahora estoy de acuerdo con X que está ocurriendo (ahora), pero ahora disiento de X, que estaba ocurriendo antaño.

Esta última proposición no podría ser verdadera, sea por incompatibilidad de las actitudes de que se

habla, sea por imposibilidad de referir X a la misma acción[9]. Por ende, el primer aserto, al ser equivalente al segundo, no podría ser verdadero. Pero. según (2) del argumento [aquí] transcrito, el primer aserto *podría* ser verdadero, pues (2) se leería:

Según D_3 y D_4, A podría estar diciendo la verdad si dijera 'X ahora está correcto, pero antaño X estuvo equivocado'.

Por tanto, (2), al ser falso, no podría seguirse de la premisa inocente (1). Mas con el colapso de (2) sobreviene el derrumbe del resto del argumento.

De igual manera, si bien el segundo argumento de Moore vale contra D_1 y D_2, con tal de que se hagan ciertas presuposiciones más bien naturales respecto de las instancias temporales, no obra contra D_3 y D_4, que específicamente excluyen tales presuposiciones. Puesto que Moore cree que su argumento posee fuerza contra *cualquier* definición en que 'correcto' o 'equivocado' se refieran solamente a las actitudes del hablante, es patente que confiere a su argumento mayor valor del que tiene.

No es mi intención defender D_3 y D_4 como aparecen, pues considero que son extraviantes, aunque por razones distintas de las de Moore, y que es probable hagan que la gente pase por alto las cuestiones centrales de la ética. Pero he querido salvaguardar esas definiciones de las objeciones de *Moore*. Al actuar así quedo en libertad —que de otra manera no poseería— para enmendar esas definiciones de manera harto simple, sin necesidad de citar las cualidades no-naturales y hacer que posean (cuanto la vaguedad del lenguaje ordinario lo permita) *un* sentido siquiera que, a mi parecer, es típicamente ético. Recurriré a esto posteriormente.

9 Supongo (como me lo permite un giro común, al menos) que el tiempo necesario para emitir una oración no basta para prevenir que los 'ahoras' se refieran todos al mismo tiempo y no basta para justificar el cambio del tiempo de 'es' a 'era'.

* [T.]

Existe una consecuencia curiosa de D_3 y D_4, que viene sugerida por el segundo argumento de Moore, y que con más razón puede dar lugar a dudas respecto del convencionalismo de esas definiciones. Si A, al hablar en el t1, dijera:

(a) X está correcto;

y al hablar en tiempo posterior, t2, afirmara:

(b) X está equivocado;

entonces esta segunda aserción *no contradiría* la palabra, pues por D_3 y D_4, (a) y (b) se convertirían en:

(aa) Yo ahora estoy de acuerdo con X, que está ocurriendo, y

(bb) Yo ahora estoy en desacuerdo con X, que estaba ocurriendo. Estas proposiciones, si A las hace respectivamente en t1 y t2, son compatibles, pues el 'ahora' de (a) no se referiría al mismo tiempo del 'ahora' de (bb). Y 'X' puede designar (como debe hacerlo si estas consideraciones han de tener algún interés) la misma acción en ambas proposiciones, puesto que el cambio de 'está ocurriendo' de (aa), al 'estaba ocurriendo' de (bb) testificaría claramente que t1, en el que (aa) fue dicho, fue anterior que t2 en que se afirmó (bb). De aquí que, pues (aa), dicho por A en t1, sería compatible con (bb) dicho por A en t2, se sigue de D_3 y D_4 que (a), dicho por A en t1, es compatible con (b), dicho por A en t2. Y *si* (a) y (b) no son compatibles en ninguna circunstancia de expresión, mientras 'correcto' y 'equivocado' se empleen en cualquier sentido ético típico, se seguiría entonces que D_3 y D_4 no conservan sentido ético típico alguno. Pero ¿es tan patente que (a) y (b), proferidas de la manera mencionada, no son compatibles? Mi 'inspección' no es tan categórica a este respecto como podría ser la de Moore; pero discutir ulteriormente sobre este punto se facilitará más después de haber tratado el tercer argumento, al que ahora nos debemos dedicar.

El tercer argumento[10] puede formularse como sigue:

(1) Si A dice 'Yo estoy de acuerdo con X' y B asevera 'Yo discrepo de X', sus dictados son compatibles lógicamente.
(2) Por ende, por D_3 y D_4[11], si A dice 'X está correcto' y B afirma 'X está equivocado', sus asertos lógicamente son compatibles.
(3) Así, de acuerdo con D_3 y D_4, si A dice 'X tiene razón' y B certifica 'X está equivocado', A y B, por lo que muestran estas proposiciones, *no* difieren en su opinión.
(4) Pero, si A dice 'X está correcto' y B asevera 'X está equivocado', entonces, en cualquier sentido típico de los términos, *difieren* en su opinión, por lo que muestran estas proposiciones.
(5) Por tanto, D_3 y D_4 no dan sentido ético típico a los términos que definen.

La crítica al tercer argumento debe atender a la inferencia de (2) a (3) y a la verdad de (4). Moore, sin duda, justificaría la inferencia de (2) a (3) mediante la suposición:

(a) Cuando A y B exponen una proposición ética, difieren en sus opiniones —por lo que muestran las

10 *Ethics*, 100 *ss.*: 'Si cuando alguien dice «Esta acción está correcta» y otro responde «No, no está correcta», cada uno no hace sino aseverar algo acerca de *sus propios* sentimientos, se sigue sin más que nunca existe realmente diferencia alguna de opinión entre ellos: nunca uno contradice realmente al otro en lo que dice. Como tampoco habría cotradicción si uno dijera «Me gusta el azúcar», y replicara otro «*No* me gusta el azúcar»... Y, de hecho, porque ello [el tipo de análisis que se está considerando] implica tal consecuencia es suficiente para condenarlo.'

11 De hecho, sólo se debería mencionar D_3, puesto que el argumento no emplea la palabra 'equivocado' que define D_4. Pero menciono D_4 simplemente porque se podría refundir el argumento con gran facilidad utilizando la palabra 'equivocado' en vez de 'correcto', sin alterar su validez o invalidez. También se podría haber hecho referencia a D_1 y a D_2, puesto que el argumento, si es válido, lo será contra cualquier definición que refiera los términos éticos exclusivamente a las propias actitudes del hablante.

proposiciones —sólo si éstas son incompatibles lógicamente.

Ahora bien, si 'A y B difieren en su opinión' se toma como otra manera de decir 'A y B tienen creencias que si las expresan verbalmente los llevan a hacer declaraciones incompatibles', entonces (a) —arriba— es verdadero. Supongamos que Moore quiere que 'difieren en sus opiniones' se entienda en este sentido y que, por ende, tiene derecho a pasar de (2) a (3), *vía* (a). En tal caso, para hacer que la proposición sea válida, debemos suponer que (4) aplica 'difieren en su opinión' en este mismo sentido. Pero la fuerza de mi crítica está en que (4), interpretado de esta manera, no es obvio en modo alguno.

Concedo que *es* obvio que, en cualquier sentido ético típico, cuando A y B aseveran 'X está correcto' y 'X está equivocado', respectivamente, difieren o discrepan en *algún* sentido. Pero no concedo que A y B, en tal caso, deben tener 'diferente opinión', en el sentido de la frase, que suponemos es el que quiere darle Moore. Creo que Moore llegó falsamente a afirmar (4) simplemente porque, debido a un énfasis exagerado en los aspectos puramente cognoscitivos del lenguaje ético, no podía entender cómo la gente podía diferir o discrepar en algún sentido, sin diferir en opinión en el sentido estrecho arriba expresado.

El sentido en que A y B, que afirman respectivamente 'X está correcto' y 'X está equivocado', 'difieren' de manera clara es un sentido que resguardaré mediante la frase 'difieren en actitud'. Y A y B diferirán en actitud cuando posean actitudes opuestas respecto de algo y cuando al menos una de ellas haga por alterar la actitud de la otra. He propugnado en otra parte que el desacuerdo en este sentido es en gran manera típico de las argumentaciones de ética; así, pues, no desarrollaré este punto aquí [12]. Bastará

12 'The Emotive Meaning of Ethical Terms', *Mind*, vol. XLVI, n. s., núm. 181. Empleo aquí 'actitud' cada vez que allí utilicé 'interés'.

con señalar que la discrepancia en la actitud con frecuencia conduce a la *discusión*, en la que cada persona expresa opiniones tales que pueden llevar al oponente, si éste las acepta, a adoptar actitud diferente al acabar la disputa. A menudo las actitudes son funciones de las creencias, por lo que no rara vez expresamos creencias en la esperanza de alterar las actitudes. Quizá Moore confundió discrepancia en la actitud con 'diferencia de opinión', y fue esta confusión la que lo llevó a aseverar (4).

Naturalmente, 'diferencia de opinión' *puede* entenderse como que significa 'discrepancia en la actitud', pero si fue esto lo que Moore entendía, no tendría derecho a pasar de (2) a (3), y el tercer argumento caería también, por más que (4) fuera verdadero.

Adviértase que cuando la gente discrepa en la actitud, no es preciso que tenga creencia falsa acerca de la propia actitud o de la ajena. Si A dice 'X está correcto' y B asevera 'X está equivocado', y uno y otro aceptan D_3, entonces es muy posible que A y B sepan *a la vez* que A acepta X y que B la rechaza. Asimismo, pueden discrepar en la actitud, pues no se están describiendo mutuamente las actitudes; para decirlo con frases de Frank Ramsey, no están 'comparando notas introspectivas'. Como tampoco están interesados en saber la *verdad* acerca de las actitudes *presentes* del otro, sino que más bien tratan de *cambiar* las actitudes del otro, en la espera de que posteriormente las actitudes de uno y otro serán del mismo tipo. No es necesario que sus juicios éticos sean incompatibles lógicamente, para que indiquen discrepancia en la actitud.

Concedido, pues, que si alguien posee sentimiento introspectivo frente a que los juicios que parecen verbalmente incompatibles respecto de lo correcto y lo equivocado son realmente incompatibles, tal sentimiento puede testimoniar sólo que existe discrepancia en la actitud y no incompatibilidad lógica. También quizá el hecho de que la gente que discrepa en la ectitud expresa con frecuencia aserciones incompa-

tibles sobre las consecuencias del objeto de la actitud, etcétera, en el curso de la discusión, puede llevar a hacer creer a uno —sin base— que los mismos juicios éticos han de ser incompatibles en cualquier sentido típico. En mi opinión, los términos éticos se emplean de hecho de manera tan vaga que la gente *no ha decidido ya si* 'X está correcto' (afirmado por A) y 'X está equivocado' (afirmado por B) se han de considerar incompatibles o no, ni es probable que los señores A y B lo hayan decidido tampoco. Así, *nosotros* podemos decidirlo según más nos guste, mientras permanezcamos fieles a las instancias que suelen suscitar las argumentaciones de ética. Bajo ciertas circunstancias de expresión, aunque no en todas, podemos convertir los juicios en incompatibles. He tratado de esto en mi escrito «Persuasive Definitions» [13], pero aquí sólo dispongo de lugar para decir que puede idearse un procedimiento que esquive las objeciones de Moore. Por otra parte, podemos hacer que los juicios expresados respectivamente por A y B se conviertan en compatibles, como se ha hecho en D_3 y D_4. Una y otra opción —por lo que me es dado ver— permiten que los términos éticos toquen las instancias que las argumentaciones de ética suelen suscitar en la vida diaria, aunque es claro que no permiten que los términos se empleen de una manera que algunos filósofos, en su confusión, desearían usarlos. No pretendo buscar para este punto certeza sobrehumana, ni me puedo extender cual sería mi deseo, pero espero haber dicho lo suficiente para demostrar que D_3 y D_4 ofrecen opciones serias frente a la cualidad no-natural de Moore.

Debo añadir, sin embargo, que D_3 y D_4 desconciertan, por cuanto no indican propiamente discrepancia en actitud; por el contrario, señalan demasiado fuertemente mera 'comparación de notas introspectivas'. Pero esto se puede remediar cualificando a D_3 y D_4 —como he hecho anteriormente— de manera harto

13 *Mind*, vol. XLVII, n. s., núm. 187.

simple. 'Correcto' y 'equivocado', y los demás términos éticos, poseen todos significado emotivo más fuerte que cualquier otro término puramente psicológico. Tal significado emotivo no se conserva en D_3 y D_4, sino que se ha de mencionar separadamente. Tiene el efecto de permitir que los juicios éticos se empleen para alterar las actitudes del oyente y, por lo tanto, se presta a argumentos que suponen disceptación respecto de la actitud. De esta manera cualificadas, D_3 y D_4 me parece que quedan inmunes a todas las objeciones de Moore.

La consideración que antes (pág. 38 y s.) daba pie a perplejidades —a saber, que 'X está correcto' dicho por A en t1 es compatible lógicamente, de acuerdo con D_3 y D_4, con 'X estaba equivocado', dicho por A en t2— ahora tiene explicación. Es claro que, en cualquier sentido típico, tales asertos son 'opuestos' de alguna manera. Mas creo que entra muy bien dentro de los límites del uso común vago afirmar que tales proposiciones, bajo las condiciones de expresión mencionadas, se pueden ver como lógicamente compatibles, como lo darían a entender D_3 y D_4, una vez cualificadas respecto de su significado emotivo. Su *parecer* incompatibles proviene del hecho de que los juicios ejercen un tipo de *influencia* emotiva diferente (o sea, que el juicio en t2 deshace el trabajo del juicio en t1). Por ejemplo, si B fue llevado por el juicio de A en t1 a concordar en su actitud con A, podría, de no haber cambiado su actitud posteriormente, encontrarse en discordancia de actitud con A en t2. Así, hablando llana aunque inteligiblemente, B podría acusar con razón a A de 'regresar' a su 'opinión' anterior. No es preciso insistir, empero, en que esta manera clara de hablar mantiene que la aserción de A en t1 era incompatible lógicamente con su aserto en t2. ¿No se puede tomar como que significa que A ha llegado a tener una actitud y a ejercer una influencia que se oponen a su primera actitud e influencia?

Se verá claro ahora que ninguno de los argumentos que he criticado es concluyente. El *método* de argumentación de Moore, tal como lo he interpretado libremente, es muy útil. Consiste en sacar consecuencias de una definición dada y luego demostrar que tales conclusiones son 'raras' en cualquier sentido *usual* de la palabra definida. Esta 'rareza' puede dar pie a la pregunta sugestiva de si tal definición propuesta es una *petitio quaestionis*. Pero si bien tal método es proficuo, puede ser mal aplicado, ora al deducir las consecuencias de la definición propuesta, ora al juzgar si tales conclusiones demuestran que la definición propuesta probablemente es una *petitio quaestionis*. Creo que Moore ha hecho una aplicación del todo equivocada del método en una u otra de sus formas.

Por más que las argumentaciones de Moore no prueban tanto como piensa (o siquiera tanto cuanto pensaba cuando escribió sus *Ethics*) no son desperdiciables en modo alguno. Espero que este repudio de mucho de sus *Principia Ethica* [14] no se interprete, por parte de críticos negligentes, como indicación que su labor en la ética no ha servido para nada. Por mucho que Moore se haya extraviado debido al lenguaje, es harto más sensible a sus atolladeros que algunos de sus opositores naturalistas; lo que se manifiesta al examinar algunos de sus argumentos. En el segundo y tercer argumento hemos visto que D_1 y D_2 no se pueden aceptar sin cualificación. Se ha de reconocer explícitamente el carácter confundente en los juicios éticos, tanto del tiempo verbal, como de la discrepancia en la actitud y del significado emotivo. Los análisis naturalistas que pasan por alto estos particulares —que ya existían al tiempo en que Moore escribía— carecen de una perspicacia que los argumentos segundo y tercero ayudan a poner de relieve.

Para evitar que se me acuse de impericia lingüísti-

[14] Ver 'Is Goodness a Quality?', en *Aristotelian Society, Supplementary Volume*, XI, 127.

ca, quiero recalcar que D_3 y D_4 requieren ulteriores cualificaciones, además de las que aquí he apuntado. Como 'correcto' y 'equivocado' son particularmente vagos y flexibles, se pueden definir según un cierto número de maneras sin salirse de ese fangoso *continuum* que denominamos 'uso ordinario'. No existe ni siquiera *una* definición que pueda abarcar su variado empleo y quizá ni siquiera bastaría *lista* alguna de definiciones, por larga que fuera. Todo lo que cabe hacer es dar definiciones 'ejemplares' (sample definitions) y esperar eludir confusiones si se llega a entender más adecuadamente (como tan frecuentemente ha recomendado I. A. Richards) la flexibilidad del lenguaje corriente.

En particular, 'correcto' y 'equivocado' cambian de significado según los diferentes contextos. Por ejemplo, si planteamos a alguien la *pregunta* '¿Está correcto X?', de ordinario no esperamos que el oyente nos diga si *nosotros* estamos acordes con X, cual inmediatamente sugerirían D_3 y D_4. Es más probable que queramos que el oyente nos diga si *él* está conforme con X y que nos influya respecto de nuestra aprobación subsiguiente. O bien, podemos desear saber qué actitudes adoptan los otros respecto de X, y así sucesivamente. O, si por principio de cuentas sabemos que el oyente está concorde con X, podemos servirnos de la pregunta '¿Está correcto X?' para insinuar que no lo está y por este medio dar a entender que estamos en desacuerdo con el oyente en actitud; desacuerdo que luego puede resultar en discusión, en la cual se pueden exponer muchas creencias que, al expresarse de determinada manera, pueden conducir, por un hecho psicológico, a la alteración de nuestra actitud o la de nuestro oponente. Y, de nuevo, si alguien 'quiere saber' si X está correcto, de ordinario no intenta sólo caracterizar sus actitudes presentes. Se decidirá a ello forzado por algún conflicto de actitudes y se esclarecerá como resultado de sus empeños por resolver el conflicto. Considerará aspectos fácticos, de precedentes, de actitud de la sociedad,

sobre la naturaleza y las consecuencias de X, etc., que puedan determinar si en adelante logrará un estado mental en que siga concordando con X o no, reprimiendo o reorientando todos los impulsos en contrario. Hay casos en que 'correcto' se emplea de un modo que varía —ligera o considerablemente— respecto de como se indica en D_3. Son unas cuantas situaciones entre las muchas que muestran que D_3 y D_4 se han de tomar sólo como definiciones 'ejemplares'.

Pero aunque no se trate más que de definiciones 'ejemplares', D_3 y D_4 —cualificadas respecto del significado emotivo— en muchos casos son paradigmas asaz interesantes. Deseo mostrar a continuación que poseen consecuencias que pueden dar razón de algunas de las propias conclusiones de Moore.

Parece muy probable, a juzgar por advertencias similares en la página 7 de *Principia Ethica,* que Moore denegaría que

'Si ahora estoy de acuerdo con X, X está correcto'

es una proposición analítica en cualquier sentido usual de las palabras. Por D_3 es analítica, y por mi parte estoy dispuesto a aceptar tal consecuencia e insistir al mismo tiempo en que D_3 es tan convencional como cualquier definición precisa de un término vago común puede serlo, *si* D_3 se cualifica con referencia al significado emotivo. Pero lo que no concedo, sin embargo, es que tal proposición sea *trivial,* a la manera como lo son la mayoría de las proposiciones analíticas. El significado emotivo de 'correcto', de la proposición de arriba, puede contribuir a inducir en el *oyente* el que apruebe X, si lo hace el hablante. Todo oyente que no desee sufrir tal influencia puede, por consiguiente, objetar contra la proposición, aunque sea analítica. Por más que sea trivial respecto de sus aspectos cognoscitivos, no lo es frente a sus repercusiones sobre la actitud, y alguien podría rehusar hacerla, como rehusaría yo, por tal motivo. Hay veces en que yo, lo mismo que otros, deseamos inducir a a los demás a que compartan nuestras actitudes, pero pocos deseamos hacerlo siempre o proceder como

si hubiésemos de esperar que el oyente concordara con nosotros en su actitud, incluso antes de que expliquemos algo sobre qué actitud, hipotéticamente, es la nuestra. He aquí la razón de que raramente se haga el aserto anterior. Esto está muy lejos de lo que concluiría Moore, pero creo que puede explicar por qué Moore, sensible conscientemente sólo a los aspectos sutiles del lenguaje, insistiría en que los juicios en cuestión, al no ser triviales, no pueden ser analíticos.

En la página 131 de su *Ethics*, Moore procede a hacer algunas advertencias notables. Menciona, con clara concordancia, a ciertos teorizantes que

> han supuesto que la cuestión sobre si una acción *está* correcta no se puede determinar cabalmente demostrando que cualquier hombre o cualquier conjunto de hombres poseen ciertos sentimientos... acerca de ella. Admitirían que los sentimientos... de los hombres podrían atañer (have a bearing) de diversas maneras a la cuestión, pero el mero hecho de que determinado hombre o conjunto de hombres tenga tal sentimiento... jamás será suficiente —dirían— *por sí mismo* para demostrar que una acción sea correcta o esté equivocada.

Estoy en completo acuerdo con esto y, de hecho, está contenido en D_3 y D_4, con tal de que tales definiciones se cualifiquen con referencia a la disceptación en actitud y al significado emotivo. Dirimir la cuestión sobre 'qué es correcto' equivale presumiblemente (en este contexto) a liquidar el desacuerdo que pueda existir entre A y B, cuando el primero sostiene 'X está correcto' y el segundo que no lo está. Tal discordancia es desavenencia de *actitud* y sólo se asentará cuando A y B vengan a tener actitudes similares. Si hubiera más gente que tomara partido por A o B, el ajuste del desacuerdo requeriría que también esta gente acabara en actitudes similares. Ahora bien, no se puede llevar a cabo tal uniformidad de actitudes con sólo señalar qué es lo que cada uno o cada conjunto acepta. Tal procedimiento podría, como dice Moore, 'atañer de diversas maneras a la cuestión', pero el conocimiento de lo que cada uno acepta *pue-*

de fracasar en conseguir la aprobación de otros. Si se ha de alterar un asentimiento mediante creencias, se deberá echar mano de toda suerte de creencias. Puede ser incluso que se deban emplear todas las ciencias, pues las creencias que servirán colectivamente para alterar las actitudes pueden ser de todas las clases concebibles y aun así no se puede garantizar que se consiga alterarlas por este procedimiento. Por esta razón es muy difícil sostener un juicio ético. Para mantener en pie los juicios éticos no basta con probar su verdad, sino que se requiere secundar, *vía* cambios en las creencias —por ejemplo— la influencia que pueden ejercer. Acepto, pues, las cuestiones anteriores de Moore, pero es patente cuán diversas son mis razones para ello.

Deseo poner en claro que, si bien el análisis a lo largo de las líneas de D_3 y D_4, respecto del significado emotivo y la discrepancia en actitud, aparece como una opción frente a las miras no-naturales de Moore, no desecha positivamente el punto de vista de que 'correcto' tiene que ver con la cualidad no-natural, sea directa sea indirectamente. Qué es lo que ahora diría Moore acerca de 'correcto', no lo sé, pero *podría* afirmar, sin rechazar el significado emotivo o la discrepancia en actitud, que '**X** está correcto' a veces significa que **X** posee alguna cualidad que es totalmente inaccesible al descubrimiento por medios científicos. Entonces 'correcto' podría tener significado emotivo, pero sólo porque designaría tal cualidad. Si se tratara de una cualidad que presupuestamente suscitara aprobación, su nombre recibiría halo laudatorio, y se vería que la gente disceptaría en actitud acerca de lo que es correcto, pero sólo porque aprobaría o desaprobaría algo según creyera o no que tal cualidad estuviera conexa con ello. Si es esto lo que Moore desea sostener y si, realmente, confía que puede hallar tal cualidad en su experiencia o 'intuición' y si, además, está seguro de que esa cualidad es no-natural, entonces no puedo aspirar a haber dicho aquí algo que pudiera convencerlo de lo contra-

rio, por más que en particular yo mantendría mis sospechas de que se ha dedicado a construir especiosamente ficciones rebuscadas, en nombre del sentido común. Arguyo, con todo, que si Moore quisiera mantener tal punto de vista, debería propugnarlo de manera más positiva, pues no puede blandirlo como la única opción para manifestar las debilidades del naturalismo. El tipo de naturalismo que combatía y que pasaba por alto la discrepancia en la actitud y el significado emotivo requiere en efecto de una alternativa, pero a menos que se hallen nuevos argumentos en contrario, tal opción sólo se puede desenvolver a lo largo de las líneas que he señalado aquí [15].

Tal alternativa, debo añadir, está muy lejos de alegar que los juicios éticos representan una 'confusión total'. El adscribir a un juicio un significado que sea en parte emotivo no quiere decir en modo alguno atribuirle confusión. Si al significado emotivo se le atribuyera algo que en realidad no fuera, entonces sin duda surgiría la confusión; pero si el significado emotivo se toma por lo que es, queda como una parte del significado inconfundible que los juicios éticos manifiestan poseer. Tampoco este tipo de análisis da a suponer peregrinamente que las instancias éticas son 'artificiales'. Las instancias que provienen de la discrepancia en la actitud, lejos de ser artificiales, son precisamente aquellas contingencias que todos nosotros necesitamos resolver de manera ineludible. No hay nadie de nosotros que esté tan ausente de la sociedad que pueda contemplar las actitudes divergentes de otros sin sentir irreprimibles ansias de tomar partido, en espera de que unas actitudes prepon-

15 Si se quieren ver análisis muy parecidos al que he llevado a cabo aquí, consúltese: A. J. Ayer, *Language, Truth and Logic*, Capítulo VI; B. Russell, *Religion and Science*, Cap. IX; W. H. F. Barnes: 'A Suggestion about Values', en *Analysis* (marzo, 1934); C. D. Broad, 'Is «Goodness» a Name of a Simple, Non-Natural Quality?', en *Proceedings of the Aristotelian Society* (1933-4) (donde se reconoce a Duncan Jones), y R. Carnap, *Philosophy and Logical Syntax*, Sec. 4.

deren sobre las demás. Nadie de nosotros es 'aislacionista' en *todos* los asuntos, puesto que todo cuanto los otros hacen o aceptan, con mucha frecuencia nos toca de cerca. Aquí, y temporalmente, me he abstenido de tomar partido en asuntos morales, pero ello ha sido de manera exclusiva por mantener mi *análisis* de los juicios morales aparte de cualquier empeño mío por ejercer influencia moral. Tal retraimiento temporal de ninguna manera implica —como apenas si vale la pena advertir— que yo considere artificiales las instancias éticas o que sostenga, para decirlo con burda paradoja, que es erróneo discutir qué está correcto o qué está equivocado.

II

REPLICA A MIS CRITICOS

G. E. Moore

ETICA

1. *¿Es 'correcto' el nombre de una característica?*

En las páginas 57 y 58 de su ensayo, el señor Broad dice que la discusión completa de mi 'doctrina' acerca de que la palabra 'bueno', cuando se emplea de una manera particular que yo tenía en mente, 'es el nombre de una característica simple y «no-natural»', debería empezar con la cuestión '¿es en efecto «bueno» el nombre de una característica?'. Naturalmente, lo que quiere dar a entender es si 'bueno', *cuando se emplea de esa manera particular*, es el nombre de una característica. Estoy de acuerdo con él en que ésta es la primera cuestión que se debería discutir, si se desea tratar por completo la 'doctrina' en cuestión.

Por su parte, él no discute esa cuestión particular ni me parece que la haya discutido ningún otro de los expositores; por tanto, tampoco la voy a discutir yo. Por fortuna, el señor Stevenson ha propuesto un

De *The Philosophy of G. E. Moore*, a cargo de P. A. Schilpp, volumen IV de Library of Living Philosophers (Northwestern University Press, Evanston, Ill., 1942), pp. 535-54. (Las futuras ediciones serán publicadas por Open Court, La Salle, Ill., y por Cambridge University Press, Londres). Reimpreso con el permiso de Library of Living Philosophers, Inc.

punto de vista acerca de los 'usos típicamente éticos' de las palabras 'correcto' y 'equivocado' que me parece evoca las mismas instancias. Si el modo de ver del señor Stevenson es atinado, creo entonces que también lo será el empleo análogo de la palabra 'bueno', que es la debatida, y se seguirá que 'bueno', según este modo de empleo, *no* es el nombre de característica alguna. Creo conveniente, por tanto, comenzar discutiendo este punto de vista del señor Stevenson.

Consideremos la proposición 'Fue correcto que Bruto apuñalara a César' o la proposición 'La acción de Bruto de apuñalar a César fue correcta' o la proposición 'Cuando Bruto apuñalaba a César estaba actuando correctamente'; tres proposiciones que parecen tener absolutamente el mismo significado. El señor Stevenson cree (p. 80) [1] que la definición '«Fue correcto que Bruto apuñalara a César» tiene el mismo significado que «Estoy de acuerdo ahora con que Bruto apuñalara a César, estaba ocurriendo»' *si se enmienda de una manera particular,* al menos *un* sentido 'típicamente ético' de esas proposiciones. Pero añade que, según cree, sólo lo hace 'mientras la vaguedad del uso ordinario lo permita. Por esta última cláusula me imagino que quiere indicar que el sentido que su definición enmendada daría a esas proposiciones sería más preciso que cualquier otro que realmente empleara alguien, si lo hiciera según el uso ordinario; pero cree que, con todo y ser más preciso, *abarca* (approaches) al menos *un* sentido en el que tal persona podría usarlas. Cree, además, que el sentido que abarca es 'típicamente ético'. Mas al afirmar que su definición enmendada ofrece (aproximadamente) al menos *un* sentido 'típicamente ético', está concediendo que por lo menos puede haber otros sentidos 'típicamente éticos' que estén igualmente acordes con el empleo ordinario, que [su defi-

1 [P. 46 y 47 de este volumen. E.]
[T.]

nición] [T.] no da ni siquiera aproximadamente. Concede, asimismo, que puede haber posiblemente otros sentidos, de igual manera concordes con el uso ordinario, que no sean 'típicamente éticos', y que tampoco da su definición ni siquiera aproximadamente. Es una generosa concesión de posibles sentidos, todos acordes con el empleo ordinario, según los cuales se podrían emplear estas sencillas proposiciones; pero a lo mejor no es tan generosa la concesión, puesto que si es circunspecta y limitada la proposición del señor Stevenson, creo que se basta para suscitar importantes cuestiones.

Parece conveniente que, antes de pasar a discutir si el señor Stevenson está acertado en su proposición circunspecta, deberíamos conocer cuál es su definición enmendada. Asevera que nos la ofrece en la página 84 [2]. Dice que tal enmienda es muy sencilla, y posiblemente lo sea, mas no es tan sencillo averiguar, por lo que dice en esa página, cuál es la enmienda de que habla. En razón de brevedad, llamaremos a la proposición 'Fue correcto que Bruto apuñalara a César' *'definiendum'* y a la proposición 'Estoy de acuerdo ahora con que Bruto apuñalara a César, que estaba ocurriendo', *'definiens'*. La definición original establecía que el *definiendum*, cuando se emplea en el sentido particular (que se aproximaba al uso ordinario) que el señor Stevenson quiere 'darnos', tiene el mismo significado que el *definiens*. Esa definición, nos dice el señor Stevenson, *tal como está* no nos da el sentido que indica, sino que debe enmendarse. Es obvio, por lo que dice, que la enmienda que se requiere tendrá algo que ver con el 'significado emotivo': o mencionará explícitamente el 'significado emotivo' o mencionará algún significado emotivo particular que una proposición pueda tener. Para ayudarnos a ver qué es la enmienda (o, como ahora la llama, la 'cualificación'), el señor Stevenson nos dice: '«Correcto», «equivocado» y otros términos éticos tie-

2 [P. 45 y 46 de este volumen. E.]

nen todos un significado emotivo más fuerte que cualquier otro término puramente psicológico'. Me imagino que con esto quiere dar a entender que el *definiendum* tiene sentido emotivo más fuerte que el *definiens*. Luego, añade: 'Este significado emotivo no se conserva en' la definición original, sino que 'debe mencionarse por separado'. Según pienso, aquí por '*debe* mencionarse por separado' quiere indicar 'debe' en la definición enmendada y en cualquier definición que nos deba 'dar' el sentido del *definiendum* que quiere darnos. Las dos proposiciones, según veo, son toda la ayuda que nos proporciona. Bien, pues, sirviéndonos de esa ayuda, ¿cuál *es* la definición enmendada? ¿Dice solamente: El *definiendum* (cuando se emplea en el sentido en cuestión) tiene el mismo significado que el *definiens*, pero posee un significado emotivo del que carece el *definiens?* ¿O dice: Posee el mismo significado, pero tiene significado emotivo *más fuerte* que el *definiens?* Si una u otra cosa es todo, sin duda no nos da sentido alguno del *definiendum* más allá del que nos da el *definiens;* sólo nos dice algo sobre un sentido posible. ¿O se trata de una proposición que mencione algún significado emotivo particular y que diga: El *definiendum* (cuando se emplea en el sentido en cuestión) tiene le mismo significado que el *definiens*, pero posee además *este* significado emotivo del que carece el *definiens?* ¿O ha de mencionar tanto algún significado emotivo particular, *como* algún grado particular de fuerza, en el que determinada frase concentre ese significado emotivo, y que diga: El *definiendum* (cuando se emplea en el sentido en cuestión) tiene el mismo significado que el *definiens*, pero posee *ese* significado emotivo en un grado de fuerza sobre *ese* grado, mientras que el *definiens* sólo lo tiene en un grado de fuerza por debajo de *ese grado?* En esos dos casos, la definición enmendada nos podría dar realmente algún sentido del *definiendum*, pero es cierto que el señor Stevenson no nos ha ofrecido ninguna enmienda de esa clase. Quizás existen otras alternativas además de estas cuatro,

pero ¿cómo vamos a poder saber qué es lo que quiere decir el señor Stevenson? El hecho escueto es que no nos ha dado *ningún* sentido del *definiendum* que vaya más allá del que nos da el *definiens*, así como tampoco ninguna definición enmendada que nos rinda tal sentido. Creo, no obstante, que es posible inferir, por lo que dice, los siguientes puntos de vista. En analogía con el modo como el señor Stevenson emplea la palabra 'cognoscitivo' y también en analogía con el uso que hace de la frase 'significado emotivo', permítasenos distinguir entre el 'significado cognoscitivo' de una oración y su 'significado emotivo'. Creo que entonces podemos decir que el señor Stevenson piensa que el *definiendum*, cuando se emplea en el sentido que tiene en mente, tiene exactamente el mismo 'significado cognoscitivo' que el *definiens*, pero, no obstante, no tiene el mismo *sentido*, porque posee 'sentido emotivo' diferente. Mas, ¿qué significa esto? ¿Cómo empleamos el término 'significado cognoscitivo'? Creo que esto se puede explicar de la manera que sigue: según el uso, algunas frases se pueden emplear de tal modo que se puede decir de quien las utilice que está *haciendo un aserto* por medio de ellas. Por ejemplo, nuestro *definiendum*, la proposición '¿Fue correcto que Bruto apuñalara a César?', puede emplearse de modo que la persona que la use esté profiriendo a todas vistas que *fue* correcto que Bruto apuñalara a César. Pero hay veces, al menos, que cuando una proposición se emplea de tal manera que la persona que la expresa está haciendo una aserción por medio de ella, está afirmando algo que concebiblemente pueda ser verdadero o falso; algo que es lógicamente posible que sea verdadero o que sea falso. Permítasenos decir que una proposición tiene 'significado cognoscitivo' si y sólo si es a la vez verdadera y se puede emplear para emitir un aserto, y también si todo aquél que la empleara en esa forma estuviera aseverando algo que podría ser verdadero o falso; y permítasenos decir también que una proposición, *p*, tiene el *mismo significado cognoscitivo*

que otra, *q*, si y sólo si tanto *p* como *q* tienen significado cognoscitivo, y asimismo que *mientras* cualquiera que empleara *p* para afirmar, aseverara algo que podría ser verdadero o que podría ser falso, habría estado afirmando exactamente lo mismo si a su vez hubiera echado mano de *q*. Si la cosa está así, el punto de vista que atribuyo al señor Stevenson es que si una persona se sirviera de nuestro *definiendum* para hacer una aserción y lo empleara en el sentido que lleva en mente el señor Stevenson, entonces, mientras afirmara algo que pudiera ser verdadero o pudiera ser falso, podría haber aseverado la misma cosa exactamente empleando en su vez el *definiens*, pero, de haber procedido así, *no* habría estado empleando el *definiens* en el mismo *sentido* en el que realmente empleó el *definiendum* y, por ende, no habría estado aseverando que fuera correcto que Bruto apuñalara a César, en el sentido que el señor Stevenson quiere. En breve, defiende el señor Stevenson que existe al menos un sentido 'típicamente ético' según el cual se puede afirmar que fue correcto que Bruto apuñalara a César, aunque la única aserción que puede ser verdadera o falsa, y que está haciendo, será que él mismo, en el momento de hablar, 'está de acuerdo en que Bruto apuñalara a César, que estaba ocurriendo'; sin embargo, del mero hecho de que está efectuando tal afirmación no se sigue que esté aseverando que la acción de Bruto fuera correcta, en el sentido en cuestión; que hace tal afirmación se seguirá de la *conjunción* del hecho de que está aseverando que 'está de acuerdo en que Bruto apuñalara a César, que estaba ocurriendo', con el hecho de que emplea palabras que poseen cierto significado emotivo (*cuál* sea el significado emotivo, no nos lo ha dicho el señor Stevenson). Existe —parece dar a entender el señor Stevenson— un tipo al menos de aserción ética tal que se distingue de otra aserción posible, que en modo alguno fuera ética, no por el hecho de que afirme algo que pueda ser verdadero o

falso y que la otra no afirme, sino simplemente por su 'significado emotivo'.

Sostiene, por tanto, el señor Stevenson —si es que lo he comprendido debidamente— que existe al menos un sentido 'típicamente ético', según el cual alguien podría aseverar que fue correcto que Bruto apuñalara a César, tal que (1) el hablante aseveraría que él, en el momento de hablar, aprobaba la acción de Bruto y (2) no estaría afirmando nada que concebiblemente pudiera ser verdadero o falso, *excepto* esto o, posiblemente también, cosas vinculadas con eso, como —*v. gr.*— que Bruto apuñaló a César. Creo, por otra parte, que tiene razón al suponer que, si bien esta proposición es limitada, no es consistente con lo que he asentado o supuesto en mis escritos sobre ética. He supuesto —creo— que *no* existe sentido 'ético típico' alguno según el cual alguien pueda afirmar que *las dos* cosas son verdaderas, y he supuesto también —creo— que no existe sentido 'ético típico' en que *ninguna* de las dos cosas sea verdadera. Diré separadamente algo acerca de estas dos propuestas separadas del señor Stevenson.

(1) Todavía me siento inclinado a pensar que no existe sentido 'ético típico' de 'Fue correcto que Bruto apuñalara a César', tal que un hombre que aseverara que fue correcto en ese sentido, *afirmara* —por regla general— que aprobaba la acción de Bruto. Opino que existe ciertamente un sentido 'ético típico', tal que quien aseverara que la acción de Bruto fue correcta en ese sentido, *supondría* * que en el momento de hablar estaba acorde con ella, o no discordaba o al menos tenía algún tipo de 'actitud' mental hacia ella. (No creo que el señor Stevenson quiera insistir en la palabra 'aprobar' porque exprese justísimamente lo

* Debo advertir que en inglés dice *imply*, pero *implicar* en español lleva el sentido de involucrar, y el *imply* de aquí se refiere a algo que no es de la esencia del afirmar. Más adelante el propio Moore da expresamente a *imply* el sentido *esencialista* y entonces sí lo he traducido por *implicar*. T.

que él quiere indicar. Creo que la esencia de esta manera de ver es sólo que existe *algún* tipo de 'actitud', tal que alguien podría aseverar, si empleara las palabras en el sentido que quiere el señor Stevenson, que tenía —en el momento de hablar— tal actitud respecto de ello). Pero creo que en todo caso, y regularmente, alguien *supondría* esto en un sentido en que decir que lo *supone* no equivaldría a decir que lo afirma y ni siquiera que se *sigue* de cualquier cosa que haya afirmado. Creo que el sentido de *suponer*, que está en cuestión, es semejante a aquél según el cual, cuando alguien asevera algo que puede ser verdadero o falso, *supone* que él mismo y al momento de hablar cree o conoce la cosa en cuestión; sentido en el que *supone* esto, aunque esté mintiendo. Si, por ejemplo, asevero un día particular que el martes anterior fui al cine, *supongo*, por el hecho de afirmar tal cosa que, en el momento de hablar, creo o sé que fui, aunque no *diga* que lo crea o lo sepa. Pero en este caso es del todo claro que esto que *supongo* no es parte de lo que *asevero*, puesto que si lo fuera, entonces para que alguien descubriera si fui al cine ese martes precisaría cerciorarse de que cuando dije que fui, yo creía o sabía que fui, lo que claramente no es el asunto de que se trata. Y también es claro que, por lo que asevero, a saber, que fui al cine ese martes, no se *sigue* que no crea o sepa que fui, cuando lo digo; pues podría haber ocurrido que hubiera ido y, no obstante, no creer o saber —en el momento de hablar— que fuera. De manera similar, pienso que si una persona sostuviera que fue correcto por parte de Bruto que apuñalara a César, aunque *supusiera* que, en el momento de hablar, aprobaba, o tenía alguna actitud semejante respecto de esa acción de Bruto, sin embargo, *no afirmaría* lo que estaba suponiendo ni se seguiría esto de nada —verdadero o falso— que estuviera aseverando. *Al decir* que la acción de Bruto fue correcta supondría que la aprobaba, pero no *estaría diciendo* que lo hiciera, ni nada de lo que dijera (si algo dijera) *implicaría* (en el senti-

do de 'conllevar' [entail]) que lo creo o sé. Creo que la seguridad aparente del señor Stevenson en que, al menos en un sentido 'típicamente ético', quien aseverara que fue correcto que Bruto apuñalara a César, estaría *aseverando* que aprobaba tal acción, puede deberse en parte a que jamás se le haya ocurrido esta alternativa de que sólo *lo diera por supuesto.* Pero creo que también se puede deber en parte a que se retrae de la paradoja que surgiría al afirmar que, incluso cuando se puede decir con toda propiedad que alguien *está aseverando* que la acción de Bruto fue correcta, con todo, podría no estar aseverando *nada en absoluto que pudiera posiblemente ser cierto o falso* —que sus palabras no tuvieran sentido *cognoscitivo* alguno— excepto, quizá, que Bruto apuñaló a César. Esta paradoja, creo, no es de mayor calibre que otras que el señor Stevenson está dispuesto a aceptar y opino que muy posiblemente pudiera ser cierta. Por lo que me es dado entender, creo que el verdadero punto de vista del señor Stevenson es que a veces, cuando alguien *afirma* que fue correcto por parte de Bruto asesinar a César, el sentido de sus palabras es (más o menos) el mismo que si dijera 'Apruebo la acción de Bruto: lo apruebo también' en que la primera cláusula daría el significado *cognoscitivo* y la segunda el *emotivo.* Pero ¿por qué no habría de decir, en cambio, que el sentido de las palabras de ese individuo es *meramente* '¡Apruebo el asesinato de César perpetrado por Bruto!', imperativo que no tuviera sentido *cognoscitivo* alguno, en el sentido que he tratado de explicar? Si esto fuera así, tal persona podría *suponer* que aprobaba la acción de Bruto, aunque no lo *dijera,* y no diría nada en absoluto que pudiera ser cierto o falso, excepto —quizá— que Bruto apuñaló a César. Ciertamente parece raro —paradójico— que pueda ser correcto decir que el hombre estaba *aseverando* que la acción de Bruto fue correcta, cuando el único significado que tendrían sus palabras sería este imperativo. ¿No podría ser éste el caso? Es más probable, a mi modo

de ver, que sea este el caso de que se trata, que sea verdadero el punto de vista del señor Stevenson.

No me parece que haya nada misterioso en este sentido de 'suponer', según el cual si afirmas que fuiste al cine el martes pasado, *supones* —aunque no *aseveras*— que crees o sabes que fuiste; y según el cual, si afirmas que la acción de Bruto fue correcta, *supones*, aunque no lo *afirmas*, que apruebas la acción de Bruto. En el primer caso, el que supongas esta proposición acerca de tu actitud presente, aunque no quede implicada por (a saber, no se siga de) *lo que* aseveras, surge simplemente del hecho, que todos aprendemos por experiencia, de que en la inmensa mayoría de los casos quien hace aserción como ésta, cree o sabe lo que asevera; el mentir, aunque sea harto común, es todavía mucho más excepcional. Por esto, decir algo como 'Fui al cine el martes pasado, pero no creo que fuera' es cosa del todo absurda, si bien *lo que* se asevera es algo que lógicamente es del todo posible: es perfectamente posible que fueras al cine y, a pesar de ello, no creyeras que hubieras asistido. La proposición donde se dice que acudiste no 'implica' que crees que fueras; que creas que acudiste no se *sigue del* hecho de que asistieras. Y, naturalmente, también del hecho de que digas que fuiste no se sigue que crees que asististe, pues podrías estar mintiendo. Pero, no obstante, el decir que acudiste *implica* (en otro caso) que crees que fuiste; por esto, decir 'Fui, pero no creo que fuera' es algo absurdo. Similarmente, el hecho de que, si afirmas que fue correcto que Bruto apuñalara a César, *supone* que apruebas o tienes tal actitud respecto de la acción de Bruto, surge simplemente del hecho, que hemos aprendido por experiencia, de que quien hace este tipo de afirmación, en la mayoría de los casos está acorde con la acción que afirma que está correcta. De aquí que, si oyéramos que alguien asevera que la acción fue correcta, presumiríamos que, a menos que estuviera mintiendo, en el momento de hablar la aprueba, aunque *no* haya aseverado que así es.

(2) Consideremos ahora la segunda parte del punto de vista del señor Stevenson; a saber, la parte donde afirma que, en algunos casos 'típicamente éticos', quien asevera que fue correcto por parte de Bruto acabar con César no asevera algo que concebiblemente pueda ser verdadero o falso, *excepto* que aprueba la acción de Bruto y también, posiblemente, que Bruto matara a César. Creo que este modo de ver las cosas es meramente negativo, pues *no* asevera que haya algunos casos en que tal hombre *esté* afirmando que concuerda con la acción de Bruto, sino que sólo afirma que existen casos en los cuales *no* afirma nada *más*, dejando del todo abierta la posibilidad de que, en todos esos casos, no esté afirmando *nada en absoluto*, que concebiblemente pudiera ser verdadero o falso. Es claro que el señor Stevenson no expresa creencia alguna respecto de que pueda haber algún caso en que tal persona, si emplea el *definiendum* en sentido 'típicamente ético' no aseverara *nada en absoluto* que concebiblemente pudiera ser verdadero o falso. Pero supone que si se consideran otras proposiciones que no sean las proposiciones (1) que aprueban la acción de Bruto, y (2) que Bruto apuñaló a César, y (3) la conjunción de las dos, entonces existirán casos en que tal hombre no aseveraría ninguna de esas *otras* proposiciones. Este es el punto de vista que quiero ahora ponderar.

Ciertamente no está de acuerdo con puntos de vista que he expresado o presumido. He supuesto, sin duda, que en todos los casos en que alguien aseverara en un sentido 'típicamente ético' que fue correcto que Bruto asesinara a César, afirmaría algo, capaz de verdad o de falsedad (o sea, alguna proposición), que a la vez (a) no sería idéntico con ninguna de las tres proposiciones citadas, (b) que no se seguiría de (3), y (c) se trataría de una proposición de la que no se seguiría (1); sería por tanto una proposición que podría haber sido verdadera, incluso si no hubiera aprobado la acción de Bruto, y que podría ser falsa incluso si la hubiera aprobado; en breve, que sería del

todo independiente lógicamente de la proposición en la que él está de acuerdo con la acción.

¿Qué hemos de decir de estos dos puntos de vista incompatibles: de la segunda parte del punto de vista de Stevenson y del punto de vista, supuesto en mis escritos, que ya he formulado?

Creo que, ante todo, debo esclarecer lo más posible cuál es mi actitud personal presente respecto de ambos. Creo ciertamente que esta segunda parte del punto de vista del señor Stevenson *puede ser* verdadera; o sea, creo sin duda que no *sé* que no sea verdadera. Mas esto no es todo. Tengo alguna propensión a pensar que *es* verdadera y que, por tanto, mi punto de vista anterior es falso. Pero al pensar, como pienso, que la primera parte del punto de vista del señor Stevenson es falsa, tengo alguna inclinación a pensar que hay al menos *un* sentido 'típicamente ético' de la proposición 'Fue correcto que Bruto apuñalara a César', tal que quien empleara esa frase en dicho sentido y la utilizara de tal manera que se pudiera decir con propiedad que estuviera *aseverando* que tal acción de Bruto fue correcta, no estaría afirmando, empero, nada en absoluto que concebiblemente pudiera ser verdadero o falso, excepto —quizá— que Bruto asesinó a César; a saber, nada *sobre* la acción de Bruto, salvo, simplemente, que acaeció. Y, yendo más allá de la precavida aserción del señor Stevenson, siento muy fuerte inclinación a pensar que, *si* existe al menos *un* sentido 'típicamente ético' según el cual son ciertas estas cosas, entonces lo son según *todos* los sentidos 'éticos típicos'. Así, pues, tengo la propensión a pensar que, en *cualquier* sentido 'típicamente ético' según el cual alguien pueda aseverar que la acción de Bruto fue correcta, no afirmaría nada en absoluto que concebiblemente pudiera ser verdadero, excepto —quizá— que ocurrió la acción de Bruto; sin ninguna particularidad, empero, como si dijera 'Por favor, cierre la puerta'. Ciertamente siento *alguna* inclinación a pensar todo esto y que, por tanto, no sólo la contradictoria, sino la contraria

de mi punto de vista anterior es verdadera. Pero, por otra parte, también siento *alguna* inclinación a pensar que mi punto de vista anterior *es* verdadero. Y si me preguntan por cuál de estos puntos de vista incompatibles siento la inclinación *más fuerte*, sólo podría responder que no sé si me siento más proclive a adherirme a uno o a otro. Creo que esto es al menos una declaración honesta de mi actitud presente.

En segundo lugar, quiero llamar la atención sobre el hecho de que, por lo que me es dado descubrir, el señor Stevenson ni da ni procura dar razón alguna que haga pensar que su enfoque es verdadero. Afirma que *puede* ser verdadero, a saber, que no sabe que lo sea, y que *cree* que lo es; pero, por lo que puedo ver, no presenta en absoluto argumentos positivos en su favor: sólo se preocupa de mostrar que ciertos argumentos que podrían usarse en contra no concluyen. Quizá *pudiera* dar algunas razones positivas que lleven a pensar que es verdadero. Pero, por lo que a mí me incumbe y aunque —como digo— siento alguna inclinación a pensar que es verdadero y si bien no sé si no poseo tanta inclinación a pensar así o a pensar que mi manera de ver anterior era cierta, no puedo dar razón alguna positiva en su favor.

Y ahora, ¿qué decir de las razones que pueda haber para pensar que el punto de vista del señor Stevenson es falso y que mi manera de ver anterior es verdadera? Puedo dar al menos una razón para ello, a saber, que *parece como si* siempre que alguien, empleando 'correcto' en sentido 'típicamente ético', asevera que una acción particular es correcta, entonces si otro, utilizando 'correcto' en el mismo sentido, asevera que no lo es, están haciendo aserciones que lógicamente son incompatibles. Si esto, que parece ser cierto, realmente lo fuera, entonces el punto de vista del señor Stevenson sería falso. Pero, realmente, del hecho de que parezca cierto no se sigue que realmente lo sea, y el señor Stevenson sugiere que parece serlo no porque lo sea sino porque, cuando ocurre

tal cosa, las dos personas —puesto el caso de que sean sinceras— *difieren* realmente *en actitud* respecto de la acción en cuestión, pero confundimos esta diferencia de actitud con el hecho de mantener opiniones lógicamente incompatibles. Dice incluso en un lugar (p. 82)[3] que cree que fui llevado con error a afirmar que aquellas dos personas realmente sostienen opiniones lógicamente incompatibles, porque no 'podía entender cómo la gente podía diferir o discrepar en algún sentido' sin mantener opiniones lógicamente incompatibles.

Creo ahora que, por lo que a esta particularidad toca, a saber, sobre cómo fui llevado a afirmar que tales dos personas mantienen opiniones lógicamente incompatibles, el señor Stevenson no ha dado en el clavo; esto sin duda alguna. Creo que, incluso cuando escribí *Principia Ethica*, era muy capaz de entender que si un miembro de una reunión, A, dice 'juguemos póker' y otro miembro, B, replica 'no, escuchemos discos', se puede decir con toda propiedad que A y B disienten. Lo que es cierto —creo— es que, al escribir *Ethics*, simplemente no se me había ocurrido que en el caso de nuestros dos hombres que afirmaran sinceramente, en un sentido 'típicamente ético' de 'correcto', y los dos en el mismo sentido, uno que la acción de Bruto fue correcta y el otro que no lo fue, que el desacuerdo entre ambos pudiera ser meramente de ese tipo. Ahora que el señor Stevenson me ha hecho parar mientes en que podría serlo, no sé a ciencia cierta si no es meramente de ese tipo, es decir, no sé con certeza si sostienen opiniones incompatibles; consiguientemente, estoy del todo de acuerdo con el señor Stevenson sobre que, cuando empleé el argumento 'Tales dos personas no pueden meramente aseverar que uno está de acuerdo con la acción de Bruto y el otro que no lo está, porque, de ser así, sus afirmaciones no serían lógicamente incompatibles', este tipo de argumentación era inconcluyente.

3 [P. 45 y 46 de este volumen. E.]

Y lo es porque no es cierto que sus asertos sean lógicamente incompatibles. Voy más allá incluso. Siento cierta inclinación a pensar que esos dos individuos *no* están haciendo aserciones incompatibles; que su desacuerdo *es* meramente una discrepancia de actitud, como el hombre que dijera 'Juguemos póker' y el que repusiera 'No, oigamos discos'. Y no sé que *no estoy tan* inclinado a pensar esto, como a pensar que están haciendo aserciones incompatibles. Pero ciertamente siento todavía *alguna* propensión a pensar que mi manera de ver anterior era verdadera y que *están* haciendo afirmaciones incompatibles. Y pienso que el mero hecho de que *parezca que así* es *es una* razón en su favor, aunque es claro que no es concluyente. Por lo que respecta al precavido punto de vista del señor Stevenson sobre que en al menos *un* caso 'típicamente ético' sólo disceptan en actitud y no están haciendo afirmaciones lógicamente incompatibles, no nos da —es claro— razón para pensar que así sea, ni veo alguna, por más que me sienta tan inclinado a pensar que así es, como a pensar que tenía razón en mi modo de pensar anterior. ¿Cómo, pues, se podrá dirimir si *están* haciendo aserciones incompatibles o no? Hay montones de casos en los que sabemos con seguridad que cierta gente *está* haciendo aserciones incompatibles, y montones de casos en los que sabemos de cierto no que no está haciendo tales manifestaciones incompatibles, como cuando alguien dice meramente 'Apruebo la acción de Bruto' y otro meramente asevera 'La desapruebo'. ¿Por qué ha de existir esta duda en el caso de las aserciones éticas? ¿Y cómo se puede disirparla?

Creo, por tanto, que el señor Stevenson no ha *mostrado* que mi punto de vista anterior estuviera equivocado, como tampoco ha *mostrado* que el argumento particular que empleé en su pro no fuera concluyente. Concuerdo con él en que no lo es, pero no ha *mostrado* que no lo sea; puesto que simplemente ha afirmado que, al menos en un caso 'típicamente ético', dos personas cualesquiera podrían *meramente* diferir

en la actitud y no sustentar opiniones incompatibles: no ha mostrado siquiera que *puedan*, es decir, que no es cierto que no difieran, tanto menos que sea cierto que *difieran*. Existe, empero, una afirmación que expresé en mis *Ethics* que ha demostrado definitivamente que se trata de una equivocación y creo que tal error es de un interés suficiente como para que sea mencionado.

Afirmé que de las dos premisas (1) sobre que cuando alguien asevera que una acción está correcta o equivocada, *solamente* está haciendo una aseveración respecto de sus sentimientos hacia ella, y que (2) a veces alguien tiene realmente frente a una acción dada un tipo de sentimiento que afirmaría que lo tenía, si dijo que estaba correcta, mientras que otro tiene realmente frente a la misma acción un tipo de sentimiento que afirmaría que lo tenía, si dijo que estaba equivocada, de estas dos premisas, repito, se *sigue* que la misma acción en algunos casos puede ser a la vez correcta y equivocada. Pero se trató de un error puro y simple; tal conclusión no se sigue de las premisas. Para poder verlo, y por qué, consideramos un caso particular. Supongamos que fuera cierto (a) que la dicción más correcta fuera de forma que se emplearan *correctamente*, esto es, de acuerdo con la mejor dicción, las palabras 'Fue erróneo que Bruto asesinara a César', si y sólo si con ellas se quisiera significar ni más ni menos que la persona, en el momento de hablar, repudiara la acción de Bruto, y que, por ende, las empleara *a la vez* correctamente *y* de tal manera que lo que por ellas quiera significar sea *verdadero*, si y sólo si en el momento de proferirlas repudia la acción. (Naturalmente, alguien puede emplear una frase de manera del todo correcta, incluso si lo que con ella da a entender es *falso*, sea porque esté mintiendo o porque se equivoque, y —similarmente— alguien puede emplear una frase de manera que lo que quiera significar con ella

sea *cierto*. incluso cuando no la emplee correctamente, como *v. gr.* cuando aplica una palabra equivocada para expresar lo que quiere decir, sea por un lapsus o porque se ha equivocado respecto a cuál es su uso correcto. Así, el emplear *correctamente* una frase —en el sentido explicado—, y el emplearla de manera que lo que por ella se quiera dar a entender sea *verdadero*, son dos cosas lógicamente independientes por completo una de otra: cualquiera de ellas puede ocurrir sin la otra.) Por brevedad, permítasenos emplear la frase 'podría decir *con verdad cabal* las palabras «Fue erróneo que Bruto asesinara a César»' de manera que signifique 'podría, si las dijera, emplearlas *a la vez* correctamente *y* de tal manera que lo que por ellas se indicara fuera verdadero'. Se seguirá entonces de la suposición antes hecha, que un hombre podría, en un tiempo dado, decir con verdad cabal las palabras 'Fue erróneo que Bruto asesinara a César', si y sólo si, en ese tiempo en cuestión, desaprobara la acción de Bruto; y del hecho de que discordara de esta acción se *seguiría* que podría decir esas palabras con verdad cabal, y del hecho de que las pudiera decir con verdad cabal se *seguiría* que desaprobaba la acción. Similarmente, presumamos que fuera cierto (b) que alguien pudiera, en un momento dado, decir con verdad cabal las palabras 'Fue correcto que Bruto apuñalara a César', si y sólo si en ese momento aprobara tal acción de Bruto. Y por fin, presumamos que fuera también cierto (c) que alguien, A, en un momento dado ha desaprobado realmente esa acción de Bruto y que ora ese mismo individuo, A, en otro momento la ha aprobado, ora otro cualquiera, B, en un momento dado la ha aprobado. La cuestión es: ¿Se sigue de (a), (b) y (c) tomadas conjuntamente que la acción deBruto de apuñalar a César fuera a la vez correcta y errada? Si tal cosa no se sigue, en este caso particular, entonces no se sigue de mis dos premisas (1) y (2) que a veces una acción sea a la vez correcta y equivocada, y cometí un error puro y simple cuando lo dije.

Ahora bien, de (a), (b) y (c) en conjunto *se sigue* que en un momento dado alguien pudiera decir con verdad cabal las palabras 'Fue erróneo que Bruto apuñalara a César' y que también en un momento dado alguien pudiera haber dicho con verdad cabal las palabras 'Fue correcto que Bruto apuñalara a César'. A primera vista parece muy natural pensar que si alguien pudo haber dicho con verdad cabal las palabras 'Fue erróneo que Bruto apuñalara a César, y dígase lo mismo del otro caso. Es del todo natural identificar la proposición 'Alguien podría haber dicho con verdad cabal *las* palabras «La acción de Bruto estuvo equivocada»', con la proposición 'Alguien podría haber dicho con verdad cabal que la acción de Bruto estuvo equivocada', y entonces preguntar: Si la acción de Bruto no estuvo equivocada, ¿cómo podría alguien decir jamás con verdad cabal que sí lo estuvo? De hecho, creo que esta última forma de proposición se emplea con harta frecuencia, pudiéndose usar con corrección, para significar lo mismo que significa la primera; y es peculiaridad de las premisas (1) y, por tanto, también de (a) que se siga de ellas que se *pudiera* usar correctamente en un caso diferente, y que, de emplearse así, entonces de 'Alguien podría haber dicho con verdad cabal que la acción de Bruto fue errada' se seguiría realmente que la acción de Bruto fue errada, aunque de 'Alguien podría decir con verdad cabal *las palabras* «La acción de Bruto fue errónea»' *no* se seguiría que lo fuera. Pero incluso, aparte de esta identificación, hay millares de casos en que de una proposición de la forma 'Alguien podría haber dicho con verdad cabal *las palabras* «p»', se sigue *p;* por ejemplo: de 'Alguien pudo haber dicho con verdad cabal las palabras «Bruto apuñaló a César»' se sigue realmente que Bruto apuñaló a César; y si no lo hubiera hecho, entonces nadie podría haber dicho tales palabras con verdad cabal. Fue, por tanto, muy natural que yo hubiera pensado que de (a) y (c) tomadas en conjunto se seguía realmente que la acción de Bruto estuviera equivo-

cada y de (b) y (c) tomadas juntas, que fuera correcta. Mas, no obstante, se trató de un error puro y simple. Lo que no logré ver fue que de (a) se sigue que de 'Alguien pudo haber dicho con verdad cabal *las palabras* «La acción de Bruto estuvo equivocada»' *no* se sigue que la acción de Bruto lo estuviera. Pues vimos que si (a) fuera verdadera, entonces 'Alguien pudo haber dicho con verdad cabal *las palabras* «La acción de Bruto estuvo equivocada»' sería equivalente de 'Alguien alguna vez ha desaprobado la acción de Bruto', mientras que —también— cualquiera que empleara las palabras 'La acción de Bruto estuvo equivocada' correctamente daría a entender por ellas, sin más, que él, en el momento de hablar, desaprobaba la acción de Bruto. De aquí que si (a) fuera verdadera, cualquiera que dijera 'Del hecho de que alguien pudo haber dicho con verdad cabal «La acción de Bruto estuvo equivocada» se sigue que la acción de Bruto lo estuvo', de emplear *correctamente* las últimas seis palabras, asentiría a la proposición de que del hecho de que alguien en un momento dado hubiera desaprobado la acción de Bruto se siguiera que él mismo, en el momento de hablar, desaprobaba la misma; lo que, es claro, resulta absurdamente falso. Si, por otra parte, no empleara correctamente las últimas seis palabras, lo que afirmara seguirse del hecho de que alguien hubiera en un momento dado repudiado la acción de Bruto no sería que la acción de Bruto fue errónea, sino algo diferente, para designar lo cual estaría empleando incorrectamente esas palabras. Por tanto, si (a) fuera cierto, no se seguiría del hecho de que alguien hubiera dicho en un momento dado con verdad cabal que 'la acción de Bruto estaba equivocada', que ésta lo estuviera. Cualquiera que lo dijese, indicaría al afirmarlo (si hablara correctamente) algo diferente de lo que otro cualquiera indicaría si lo dijera; y cada una de estas diferentes cosas serían absurdamente falsas. De aquí que fuera

* El original dice cuatro palabras: (Brutos' action was wrong).

un error puro y simple inferir que pues de (a) y (c) en conjunto se seguiría que alguien pudiera haber dicho con verdad cabal *las palabras* 'La acción de Bruto estuvo equivocada', por lo mismo se seguiría que la acción de Bruto estuvo equivocada: esto último no se seguiría, aunque sí lo primero. Si, por otra parte, en vez de la proposición 'Alguno podría haber dicho con verdad cabal *las palabras* «La acción de Bruto estuvo equivocada»', consideramos la proposición que contrasté antes con ésta, a saber 'Alguno podría haber dicho con verdad cabal que la acción de Bruto estuvo equivocada' este último individuo, de ser cierta (a), *podría indicar* —si fuera yo quien lo dijera— 'Alguno podría haber dicho con verdad cabal que yo ahora desecho la acción de Bruto', de lo que, es claro, se seguiría que ahora desapruebo la acción de Bruto.

Quizá se podría haber dicho todo esto harto más sencillamente; incluso así lo ha hecho el señor Stevenson. Pero en todo caso estoy del todo conforme con él en que fue un error sin más, por mi parte, afirmar que de las premisas (1) y (2) se seguiría que la misma acción era a la vez correcta y equivocada. Ha sido él quien me ha convencido de que era un error.

Quizá debería, por fin, explicar por qué dije antes que si el punto de vista del señor Stevenson sobre los usos 'típicamente éticos' de la palbra 'correcto' estuviera en lo cierto, entonces 'correcto' —empleado en sentido típicamente ético— no sería 'el nombre de una característica', y que si 'correcto' no fuera tal, tampoco lo sería 'bueno' en el sentido en el que principalmente me ocupé.

Naturalmente, no es del todo cierto que esto se siga de la manera de ver del señor Stevenson. Como he señalado, se limita cautamente a decir que, al menos en *un* sentido típicamente ético, 'correcto' se emplea de una manera particular que deja abierta la posibilidad de que, si se emplea en tal sentido, incluso si no fuera 'el nombre de una característica', con todo podrían existir otros usos éticos en los que igual-

mente fuera el nombre de una característica. Pero me parece que si existe siquiera *un* uso ético, cual sostiene el señor Stevenson que ha de existir, entonces —probablemente— *todos* los usos éticos serán equivalentes a ése en el sentido que me hace decir que, si se emplea como el señor Stevenson piensa que alguna vez ocurre, cuando entonces no sería 'el nombre de una característica'.

¿Por qué, pues, dije que 'correcto', empleado del modo que describe el señor Stevenson, no sería 'el nombre de una característica'? Temo que no pueda presentar mejor razón que ésta. Si 'correcto' se empleara de la manera en cuestión, se seguiría tanto (1) que ninguna de dos personas que, empleándolo de esa manera, dijeran de la misma acción que estaba correcta o que podría estarlo, jamás dirían la misma cosa al respecto, puesto que una diría que ella, al momento de hablar, la aprobaba, mientras que la otra diría que la *aprobaba*, como también (2) que ninguna persona en particular que en dos ocasiones distintas dijera de la misma acción que estaba correcta o podría estarlo, diría jamás la misma cosa al respecto en una ocasión y en otra, puesto que en una ocasión diría que la aprobaba en *esa* época, y en la otra circunstancia diría que la aprobaba en ese momento particular. En breve, 'correcto', si se emplea a la manera del señor Stevenson, significará cosas distintas cada vez que se emplee como predicado. Y me pareció, y aún me lo parece, que decir de una palabra que en un sentido particular es 'el nombre de una característica' se entendería naturalmente que quiere significar que, al emplearse de tal forma, significa lo mismo tanto al usarse en momentos diferentes cuando son diversas personas las que la utilizan. Si no es así, entonces no existe esa característica de la que es nombre. Es claro que se puede decir que 'correcto', empleado a la manera descrita por el señor Stevenson, sería el nombre de una y sólo de

una 'característica' cada vez que se empleara, aunque fuera diversa cada vez; si bien esto debería cualificarse diciendo que en cada ocasión, aunque fuera el nombre de una característica, no sería *meramente* el nombre de una característica, ya que también poseería 'significado emotivo'. Creo que esto estaría de acuerdo con la manera como los filósofos emplean el término 'característica' (y —pienso— en la manera como la empleaba el señor Broad), pues a veces la emplean de modo que si ahora digo 'Apruebo que Bruto apuñalara a César' esté atribuyendo a esa acción de Bruto cierta 'característica', a saber, la de ser aprobada por mí ahora. Sin duda, se trata de un empleo de la palabra 'característica' que difiere de cualquier otro corriente. Nadie tendría la ocurrencia de decir, en conversación ordinaria, que dicha acción de Bruto posee, si la apruebo ahora, una característica de que carecería si no la aprobara. De ordinario empleamos 'característica' de tal manera que 'se apruebe ahora por mí' o 'se profiera por mí ahora' no son 'características' de tal acción en modo alguno. No obstante, creo que existe un uso filosófico bien establecido en que tenga cabida, con tal de que hable ahora de tal acción o la apruebe ahora. Imagino que el señor Broad empleó 'característica' en esta modalidad filosófica. Se debe admitir, pues, que 'correcto', empleado cual describe el señor Stevenson, sería, en tal sentido de 'característica', el nombre de una característica cada vez que se empleara, aunque diferente cada vez y aunque no sería *meramente* el nombre de una característica, ya que poseería también 'significado emotivo'. Pero este hecho, según me parece, no nos justificaría decir que, de acuerdo con este empleo, era el nombre de una característica, puesto que esta última frase se entendería que naturalmente quiere decir —según ese empleo— que era el nombre de *una y misma característica cuando se empleara en diferentes momentos y por distintas personas.*

Pero decir que 'correcto', en sus usos éticos, no es 'el nombre de una característica' podría significar

también algo más que, pienso, tenía en mente el señor Broad con toda probabilidad cuando dijo que era cuestionable si 'bueno' (en ese uso particular) era en absoluto el nombre de una característica. Supongamos que, por cuanto respecta, al menos, a *uno* de los usos 'típicamente éticos' de 'correcto', lo que arriba denominé primera parte del punto de vista del señor Stevenson, fuera falso, mientras que el segundo fuera cierto, de modo que 'Fue correcto que Bruto apuñalara a César', al emplearse de esta manera, no tuviera significado *cognoscitivo* alguno (excepto, quizá, que Bruto apuñaló a César), pero fuera meramente equivalente a algún imperativo o impetración' ¡Aprueba la acción de Bruto de matar a César!', entonces, en este caso, 'correcto' —así empleado— no sería el nombre de una característica, en el sentido de que una persona que en tal sentido aseverara que fue correcto que Bruto apuñalara a César, no aseveraría nada en absoluto que pudiera ser posiblemente cierto o falso, excepto quizá simplemente que Bruto asesinó a César: al aseverar que la acción de Bruto estuvo correcta, no afirmaría nada en absoluto *respecto de* la acción, salvo el que acaeciera. Que 'correcto', en este sentido, no es el nombre de una característica, es naturalmente punto de vista que no puede ser atribuido al señor Stevenson, puesto que él sostiene sólo que la segunda parte de su tesis es cierta en casos en que también lo sea la primera; a saber, donde 'Fue correcto que Bruto apuñalara a César' tiene *algún* sentido cognoscitivo cada vez que es proferido, aunque cada vez diferente según sea el momento y la persona. Pero he dicho arriba que pensaba que era más probable que la segunda parte de este punto de vista fuera verdadera, y falsa la primera, y no que las dos juntas fueran verdaderas. Si esto fuera así, entonces 'correcto', en este sentido más radical, no sería 'el nombre de una característica'.

Debo decir, de nuevo, que me siento inclinado a pensar que 'correcto', en todos los usos éticos, y —por supuesto— 'equivocado', 'debe', 'deber' no son

tampoco —en este sentido más radical— nombres de características en modo alguno, antes bien, que sólo poseen 'significado emotivo' y ningún 'significado cognoscitivo'. Y, si esto es cierto en su caso, debe serlo también en 'bueno', a tenor del sentido que principalmente me ha ocupado. Me siento *inclinado* a pensar que así es la cosa, pero también me inclino a pensar que no lo es, sin saber hacia qué lado me inclino más. Si estas palabras, en sus usos éticos, sólo tienen significado emotivo, o si el punto de vista del señor Stevenson al respecto es verdadero, entonces ha de parecer que todo lo demás que yo vaya a decir sobre el asunto tiene que ser o fútil o falso (nonsense or false), no sé cuál de las dos cosas. Pero a mí no me parece que lo que voy a decir sea ni fútil ni falso. Ello es —creo— una razón adicional (aunque, naturalmente, no concluyente) para suponer tanto que tienen significado 'cognoscitivo', como que el punto de vista del señor Stevenson respecto de la naturaleza de este significado cognoscitivo es falso.

III

LA FALACIA NATURALISTA

W. K. Frankena

De *Mind*, vol. 48 (1939), pp. 464-77. Reimpreso con la venia del autor y del editor de *Mind*.

El historiador futuro del 'pensamiento y expresión' del siglo xx registrará sin duda con algo de diversión el prurito de algunos de los filósofos controversistas del primer cuarto de siglo en rotular los puntos de vista de sus opositores como 'falacias'. Es posible que llamen su atención algunas de estas falacias, un tanto sonoras, aplicadas por sus inventores cual títulos: la falacia de la predicación inicial, la falacia de la localización simple, la falacia de lo concreto mal situado, la falacia naturalista.

De estas falacias, reales o supuestas, la más famosa es quizá la falacia naturalista. Los factores de cierto tipo de teoría ética, predominante en Inglaterra y bien representada en América, que recibe los distintos nombres de objetivismo, no-naturalismo o intuicionismo, con frecuencia han acusado a sus impugnadores de cometer la falacia naturalista. Alguno de éstos han repudiado ásperamente el cargo de tal falacia, mientras que otros han comentado el asunto por lo menos de pasada, pero en general la noción de falacia naturalista tiene considerable circulación en la literatura ética. Con todo, a pesar de su renombre, la falacia naturalista jamás se ha discutido largo y ten-

dido, y por esta razón me he decidido a realizar un estudio de ella en este artículo. De paso espero esclarecer ciertas confusiones que se han suscitado en conexión con la falacia, pero mi interés principal es liberar la controversia entre intuicionistas y opositores de la noción de que exista la falacia lógica o cuasilógica, e indicar dónde realmente se halla el punto decisivo.

El relieve obtenido por el concepto de falacia naturalista en la filosofía moral reciente es otro testimonio de la gran influencia del filósofo de Cambridge, el señor G. E. Moore, y de su libro *Principia Ethica*. Así, el señor Taylor se refiere al 'error vulgar', que el señor Moore nos ha enseñado, consistente en hablar de la 'falacia naturalista' [1], y G. S. Jury, como para ilustrar cuán bien hemos aprendido esa lección, dice con referencia a las definiciones naturalistas de valor: 'Todas esas definiciones tienen la imputación de «falacia naturalista» [2] del Dr. Moore'. Ahora bien, el señor Moore acuñó la noción de falacia naturalista en su polémica contra los sistemas naturalistas y metafísicos de ética. 'La falacia naturalista es una falacia', escribe, y 'no debe cometerse'. Sin embargo, todas las teorías naturalistas y metafísicas de ética 'se *basan* en la falacia naturalista, en el sentido de que la comisión de esta falacia ha sido la causa principal de su amplia aceptación' [3]. La mejor manera de librarse de ellas es, pues, exponerlas a la luz. Con todo, aún no se aclara cuál es el *status* de la falacia naturalista en la polémica de los intuicionistas contra otras teorías. A veces se emplea como arma, como cuando el señor Clarke dice que si llamamos buena a una cosa simplemente porque agrada, somos culpables de falacia naturalista [4]. En efecto, en muchas partes de *Principia Ethica* se presenta también este aspecto

1 A. E. Taylor, *The Faith of a Moralist*, vol. i, p. 104 n.
2 *Value and Ethical Objectivity*, p. 58.
3 *Principia Ethica*, pp. 38, 64.
4 M. E. Clarke, 'Cognition and Affection in the Experience of Value', *Journal of Philosophy* (1938).

al lector. Ahora bien, al usarla como arma, los intuicionistas se sirven de la falacia naturalista como si fuera una falacia lógica coincidente por los cuatro costados con la falacia de composición, cuyo descubrimiento acaba con la ética naturalista y metafísica y deja campeando el intuicionismo. O sea, que se toma por adelantado, como falacia, para blandirla en la controversia. Mas existen señales en *Principia Ethica* indicadoras de que la falacia naturalista posee lugar más bien diferente en el esquema de los intuicionistas y en modo alguno debería emplearse como arma. En este aspecto se ha de probar que la falacia naturalista lo sea. No se puede emplear para dirimir la controversia, sino que se podrá confirmar que es falacia una vez que haya escampado el humo de la batalla. Consideremos los siguientes pasajes: (*a*) 'la falacia naturalista consiste en la opinión de que bueno no *significa* nada, sino una noción simple o compleja, definible por cualidades naturales'; (*b*) 'el aserto de que el bien es indefinible y que negarlo implica falacia es afirmación sometible a prueba estricta'[5]. Estos pasajes parecen suponer que la falacidad de la falacia naturalista es precisamente el quid de la controversia entre intuicionistas y contraponedores y no puede ser manejada como arma en dicha controversia. Una de las cuestiones que deseo esclarecer en este escrito es que el cargo de comisión de falacia naturalista cabe, en todo caso, sólo como conclusión del debate y no como instrumento para dirimirlo.

La noción de la falacia naturalista se ha relacionado con la noción de la bifurcación entre el 'debe' y el 'es', entre valor y hecho, entre lo normativo y lo descriptivo. Así, el señor D. C. Williams dice que algunos moralistas han pensado que es apropiado incusar como falacia naturalista el intento de derivar Debe de Es[6]. Podemos empezar, pues, considerando esta

5 *Principia Ethica*, pp. 73, 77. Ver también p. xix.

6 'Ethics as Pure Postulate', *Philosophical Review* (1933). Ver también T. Whittaker, *The Theory of Abstract Ethics*, pp. 19 *s.*

bifurcación, la que Sidgwick, Sorley y otros pusieron de relieve como reacción principalmente a los procedimientos de Mill y Spencer. Afirma Hume que esa bifurcación se halla en su *Treatise* (Tratado): 'No puedo pasar por alto añadir a estos razonamientos una observación que, quizá, sea de importancia. En todo sistema de moral que hasta ahora he examinado he advertido siempre que el autor procede durante un lapso de tiempo según la manera ordinaria de raciocinar, probando la existencia de Dios o haciendo observaciones sobre las cosas humanas; pero de repente me sorprende hallar que en vez de las cópulas ordinarias de las proposiciones —*es* y *no es*— me encuentro con que no aparece proposición que no esté conexa con un *debe* o un *no debe*. Este cambio es imperceptible, mas no obstante es de suma importancia hasta el final. Al expresar este *debe* o *no debe* algún tipo nuevo de relación o afirmación, es preciso que se observe y explique, a la par que se dé alguna razón de lo que parece del todo inconcebible, a saber, cómo esta nueva relación puede ser deducida de otras que son por entero diferentes de ella. Pero como de ordinario los autores no hacen uso de esta precaución, me permito advertírselo a los lectores. Estoy convencido de que si se parara mientes en este punto nimio, los sistemas de moral corriente sufrirían subversión, y veríamos que la diferencia entre vicio y virtud no está fundada exclusivamente en relaciones de objetos ni se percibe por la razón'[7].

Huelga decir que los intuicionistas *han* visto que esta observación es de alguna importancia[8]. Están acordes con Hume en que trastorna todos los sistemas corrientes de moral, aunque —es claro— niegan que nos permita ver que la distinción de virtud y vicio no está fundada en relaciones de objetos y que no se percibe por la razón. De hecho, sostienen que

7 Libro III, parte ii, sección i.

8 Ver J. Laird, *A Study in Moral Theory*, pp. 16 *s.*; Whittaker, *op. cit.*, p. 19.

si se para la debida atención subvierte también el propio sistema de Hume, puesto que [dicho sistema] * trae definiciones naturalistas de virtud y vicio, de bien y mal [9].

La tesis de Hume es que las conclusiones éticas no se pueden deducir válidamente de premisas que son no-éticas. Pero cuando los intuicionistas sostienen la bifurcación del 'debe' y del 'es', apuntan a algo más que a que las proposiciones éticas no se pueden deducir de proposiciones no-éticas, pues esta dificultad podría remediarse en los sistemas corrientes de moral —como veremos— introduciendo definiciones de nociones éticas en términos no-éticos. Pero sostienen, además, que son imposibles las definiciones de nociones éticas en términos no-éticos. 'El punto esencial', dice el señor Laird, 'es que los valores son irreducibles a no-valores' [10]. Pero aún sostienen más. Lo amarillo y lo placentero son, según el señor Moore, indefinibles en términos no-éticos, pero son cualidades naturales y pertenecen a la circunscripción del 'es'. Mas las propiedades no son para él meras cualidades naturales indefinibles, descriptivas o expositorias; son propiedades de *tipo* diferente, no descriptibles o no-naturales [11]. La bifurcación de los intuicionistas contiene tres proposiciones:

(1) Las proposiciones éticas no se pueden deducir de las no-éticas [12].
(2) Las características éticas no se pueden definir en términos de las no-éticas.
(3) Las características éticas son diferentes, en tipo, de las no-éticas.

En realidad sólo se trata de una proposición, de la (3), puesto que la (3) contiene la (2) y la (2) contiene la (1). Esto no quiere decir que toda característica ética sea indefinible absolutamente. Esta es otra cuestión, aunque no siempre se advierta así.

9 Ver C. D. Broad, *Five Types of Ethical Theory*, c. iv.
10 *A Study in Moral Theory*, p. 94 n.
11 Ver *Philosophical Studies*, pp. 259, 273 *s.*
12 Ver J. Laird, *op. cit.*, p. 318. También pp. 12 *ss.*

Ahora bien, ¿qué tiene que ver la falacia naturalista con la bifurcación de 'debe' y de 'es'? Para empezar, la conexión es ésta: muchos moralistas naturalistas y metafísicos proceden como si las conclusiones éticas se pudieran deducir de premisas todas las cuales fueran no-éticas, siendo clásicos ejemplos Mill y Spencer. O sea, que violan (1). Este procedimiento posteriormente ha recibido el nombre de 'falacia factualista', dado por el señor Wheelwright, y el de 'falacia valuatoria', que le ha adscrito el señor Wood [13]. El señor Moore parece a veces identificarlo con la falacia naturalista, pero en conjunto sólo sostiene que supone, implica o estriba en esta falacia [14]. Ahora podemos considerar el cargo de que el procedimiento en cuestión es o implica una falacia.

Por principio de cuentas podemos dejar señalado que, incluso si la deducción de conclusiones éticas de premisas no-éticas no es falacia en modo alguno, Mill de todas maneras la cometió al extraer una analogía entre la visibilidad y la desiderabilidad en su argumentación sobre el hedonismo, y quizá la comisión de *esta* falacia por su parte, la que —como dice el señor Broad— aprendemos ya en las rodillas de nuestras madres, es la principal promotora de la noción de la *falacia* naturalista. ¿Pero es falacia deducir conclusiones éticas de premisas no-éticas? Consideremos el argumento epicúreo sobre el hedonismo que Mill trató de embellecer tan desatinadamente: el placer es bueno, puesto que todos los hombres lo buscan. Aquí se deriva una conclusión ética de una premisa no-ética. Y, en efecto, tal argumento, cual aparece estrictamente, *es* falaz. Pero no lo es porque ocurra en la conclusión un término *ético* que no aparece en la premisa, sino que es falaz porque todo

13 P. E. Wheelwright, *A Critical Introduction to Ethics*, pp. 40-51, 91 *s.*; L. Wood, 'Cognition and Moral Value', *Journal of Philosophy* (1937), p. 237.

14 Ver *Principia Ethica*, pp. 114, 57, 43, 49. Whittaker la identifica con la falacia naturalista y la considera como falacia 'lógica', *op. cit.*, pp. 19 *s.*

argumento de la forma 'A es B, por tanto A es C' no es válido, si se toma estrictamente como aparece. Por ejemplo, no es valedero sostener que Creso es rico porque es opulento. Pero tales argumentos no se proponen para que se tomen cual aparecen. Son entimemas y contienen una premisa elidida. Cuando esta premisa elidida se hace explícita, se convierten en válidos y ya no contienen falacia lógica [15]. Así la inferencia epicúrea del hedonismo psicológico al ético es válida cuando se explicita la premisa suprimida, de manera que resulte que lo que todos los hombres buscan es el bien. Entonces lo único que queda por resolver es si las premisas son verdaderas.

Es claro, entonces, que la falacia naturalista no es una falacia lógica, puesto que puede aparecer (be involved) incluso cuando el argumento es válido. ¿Cómo se inmiscuye la falacia naturalista en tales 'argumentos éticos mixtos' [16] como el de los epicúreos? El que se inmiscuya o deje de hacerlo dependerá de la naturaleza de la premisa elidida. Esta puede ser una inducción. Si es una de las tres primeras cosas, no ocurrirá en modo alguno la falacia naturalista. De hecho, entonces el argumento no contiene violación de (1), puesto que una de las premisas será ética. Pero si la premisa que se ha de explicitar es una definición, o una proposición que es verdadera por definición, como lo era probablemente para los epicúreos, entonces el argumento, sin dejar de ser válido, contiene la falacia naturalista y será de este tipo:

(*a*) Todos los hombres buscan el placer.
(b) Lo que todos los hombres buscan es el bien (por definición).
(*c*) Luego el placer es bueno.

No me interesa sobremanera determinar si este argumento, cual aquí lo he explanado, viola (1). Si no

15 Ver *ibid.*, pp. 50, 139; Wheelwright, *loc. cit.*
16 Ver C. D. Broad, *The Mind and its Place in Nature*, pp. 488 *s.*; Laird, *loc. cit.*

lo hace, entonces ningún 'argumento ético mixto' comete realmente falacia alguna factualista o valuatoria, excepto cuando indebidamente se toma como completo en su forma entimemática. Si viola (1), entonces un argumento válido puede incluir la deducción de una conclusión ética de premisas no-éticas y la falacia factualista o valuatoria no será realmente una falacia. El quid estará en si (*b*) y (*c*) se toman como proposiciones éticas o no. El señor Moore se rehúsa a considerarlas tales, contendiendo que —por hipótesis— (*b*) es analítica o tautológica, y (*c*) es psicológica, puesto que realmente sólo dice que todos los hombres buscan el placer [17]. Mas decir que (*b*) es analítica y no-ética y que (*c*) no es ética sino psicológica, es prejuzgar la cuestión de si se puede definir el 'bien'. Pues los epicúreos sostendrían precisamente que si su definición es correcta, entonces (*b*) es ética pero analítica y (*c*) ética aunque psicológica. Así, a menos que se quiera convertir en *petitio quaestionis* la definibilidad de bondad, se habrá de considerar a (*b*) y a (*c*) como éticas, en el cual caso nuestro argumento no viola (1). Supongamos, empero, si no carece de sentido, que (*b*) es no-ética y que (*c*) es ética; entonces el argumento violará (1), pero no obstante seguirá obedeciendo a todos los cánones de la lógica, por lo que sólo sirve para confundir hablar de 'lógica valuatoria', cuya regla básica establece que no cabe deducir una conclusión valuatoria de premisas no-valuatorias [18].

La única forma como, ya los intuicionistas, ya los postulacionistas como el señor Wood, pueden echar sombras de duda sobre la conclusión del argumento de los epicúreos (o sobre la conclusión de cualquier argumento paralelo) es atacando las premisas, en particular (*b*). Ahora, según el señor Moore, si el argumento contiene la falacia naturalista, es debido a la presencia de (*b*). Implica (*b*) la identificación de bon-

17 Ver *op. cit.*, pp. 11 *s.*; 19, 38, 73, 139.
18 Ver L. Wood, *loc. cit.*

dad con 'todos los hombres buscan', pero hacer ésta o identificaciones parecidas es cometer la falacia naturalista. La falacia naturalista no es el procedimiento de violar (1), sino que es el procedimiento, supuesto en muchos argumentos éticos mixtos, y explícitamente inferido por muchos moralistas, independientemente de estos argumentos, de definir características tales como la bondad o de sustituir alguna otra característica por ellas. Bastará con citar algunos pasajes de *Principia Ethica:*

(*a*) '... han sido demasiados los filósofos que han pensado que cuando citaron esas otras propiedades [propias de todas las cosas que son buenas] realmente estaban definiendo el bien; o sea, que esas propiedades, de hecho, no eran simplemente «otras», sino absoluta y enteramente lo mismo que la bondad. A esta manera de ver las cosas propongo que se la denomine «falacia naturalista»...' [19]

(*b*) 'Así, pues, he apropiado el nombre de Naturalismo a un método particular de enfocar la ética... Tal método consiste en sustituir alguna propiedad de un objeto natural o de un conjunto de objetos naturales para que haga las veces de «bueno»...' [20]

(*c*) '... La falacia naturalista es aquélla que consiste en identificar la noción simple que indicamos por «bueno» con otra noción.' [21]

Así, identificar 'mejor' y 'más evolucionado', 'bueno' y 'deseado', etc., equivale a cometer la falacia naturalista [22]. Pero, ¿por qué exactamente tal procedimiento resulta falaz o erróneo? ¿Y se trata sólo de una falacia cuando se aplica a bueno? Ahora debemos estudiar la Sección 12 de *Principia Ethica.* Aquí, el señor Moore hace algunas aserciones interesantes:

'... si alguien quisiera definirnos lo que es el placer como si se tratara de cualquier objeto natural; si al-

19 P. 10.
20 P. 40.
21 P. 58, *cf.* pp. xiii, 73.
22 *Cf.* pp. 49, 53, 108, 139.

guien dijera, por ejemplo, que placer *significa* la sensación de rojo... Bien, entonces se trataría de la misma falacia que he llamado falacia naturalista... No debería llamarla falacia naturalista en realidad, aunque se trate de la misma falacia que he llamado naturalista con referencia a la ética... Cuando alguien confunde dos objetos naturales entre sí y define el uno por el otro... entonces no existe razón para llamar a tal falacia naturalista. Pero si confunde «bueno», que no es... un objeto natural, con otro objeto natural cualquiera, entonces hay razón en denominar a esto falacia naturalista...' [23]

Aquí, el señor Moore debería haber añadido que, cuando alguien confunde 'bueno', que no es ni un objeto ni una cualidad metafísicos, con cualquier cualidad u objeto metafísicos, como hacen los moralistas metafísicos, según él, entonces esa falacia debería recibir el nombre de metafísica. En cambio, la llama naturalista también en este caso, aunque reconoce que se trata de un caso diferente, puesto que las propiedades metafísicas son no-naturales [24], procedimiento que ha extraviado a muchos lectores de *Principia Ethica.* Por ejemplo, ha conducido al señor Broad a hablar de 'naturalismo teológico' [25].

Resumiendo: 'Incluso si [bondad] fuera un objeto natural, ello no alteraría la naturaleza de la falacia ni disminuiría en un ápice su importancia' [26].

Se ve claramente por estos pasajes que la falacia de procedimiento, que el señor Moore llama falacia naturalista, no se debe al hecho de que se aplique a bueno o a una característica ética o no-natural. Cuando el señor R. B. Perry define 'bueno' como algo que 'es objeto de interés', la dificultad no está solamente en que está definiendo *bueno,* ni en que define una característica *ética* en términos de las *no-éticas,* ni

23 P. 13.
24 Ver pp. 38-40, 110-112.
25 *Five Types of Ethical Theory,* p. 259.
26 P. 14.

en que considera una característica *no-natural* como si fuera *natural*. Se trata de un inconveniente más genérico que todo esto. Por razón de claridad hablaré de falacia definista, cual si fuera una falacia subyacente en la falacia naturalista. Entonces, según los pasajes anteriores, la falacia naturalista será una especie o forma de la falacia definista, como también lo sería la falacia metafísica, si el señor Moore hubiera dado distinto nombre a ésta [27]. Es decir, la falacia naturalista —según se ve por el procedimiento de Perry— es tal no porque sea naturalista o confunda una cualidad no-natural con alguna natural, sino solamente porque conlleva la falacia definista. Así, pues, podemos dirigir nuestra atención enteramente al entendimiento y valoración de la falacia definista.

A juzgar por los pasajes que he citado, la falacia definista es el proceso de confundir o identificar dos propiedades, de definir una propiedad por otra o de sustituir una propiedad por otra. Además, hay tal falacia siempre que dos propiedades se traten simplemente como si fueran una; no importa —si tal caso se diera— que una de ellas fuera natural o no-ética y la otra no-natural o ética. Se puede cometer la falacia definista sin incurrir en la bifurcación de lo ético y lo no-ético, como cuando se identifica el placer y lo rojo o lo correcto y lo bueno. Incluso cuando se incurre en esa bifurcación al cometer la falacia definista, como cuando se identifica lo bueno y lo placentero y la satisfacción, entonces el *error* no está en que se incurre en la bifurcación, sino en que las dos propiedades se tratan cual si fueran una. Por tanto, según esta interpretación, la *falacia* definista no consiste —en ninguna de sus formas— en violar (3), y no tiene conexión esencial alguna con la bifurcación de 'debe' y de 'es'.

Esta formulación de la falacia definista explica o refleja el lema de *Principia Ethica* tomado del obispo Butler: 'Everything is what it is, and not another

27 Como lo ha hecho Whittaker, *loc. cit.*

thing' (Todo es lo que es y no otra cosa). Se sigue de este lema que la bondad es lo que es y no otra cosa. Se sigue que los puntos de vista que intentan identificarla con algo más cometen un error de un tipo elemental. Pues *es* un error confundir o identificar dos propiedades. Si las propiedades son dos, entonces sencillamente no son idénticas. Pero, ¿cometen este error quienes definen las nociones éticas en términos no-éticos? Replicarán al señor Moore que no identifican dos propiedades; lo que están diciendo es que dos palabras o conjuntos de palabras hacen las veces o significan una e idéntica propiedad. En parte, el señor Moore fue desorientado por la forma de hablar material, como la llama el señor Carnap, en frases como 'La bondad es placer', 'El conocimiento es creencia verdadera', etc. Cuando, en cambio, alguien dice: 'La palabra «bueno» y la palabra «placentero» significan la misma cosa', etc., se ve claro que no se están identificando dos cosas. Pero el señor Moore no logró ver esto, al negar que se interesara en proposición alguna acerca del empleo de las palabras [23].

La falacia definista, pues, tal cual la hemos planteado, no excluye ninguna definición naturalista o metafísica de los términos éticos. La bondad no se puede identificar con ninguna 'otra' característica (si es que es alguna característica en absoluto). Pero el problema es éste: *¿qué* características hay, diferentes de la bondad? Es una *petitio quaestionis* decir sin más que el señor Perry, pongamos por caso, identifica la bondad con alguna otra cosa. Lo esencial es que la bondad es lo que es, aunque sea definible. Y por lo mismo, el señor Perry puede tomar como lema de su *Moral Economy* naturalista otra frase del obispo Butler: 'Things and actions are what they are, and the consequences of them will be what they will be; why then should we desire to be deceived?' (Las cosas y las acciones son lo que son y sus consecuencias

[23] Ver *op. cit.*, pp. 6, 8, 12.

serán lo que serán, ¿para qué hemos de desear que se nos engañe?) El lema de *Principia Ethica* es una tautología y debe explicarse de la siguiente manera: Cada cosa es lo que es y no otra cosa, a menos que sea otra cosa, pero aun entonces es lo que es.

Por otra parte, si el lema del señor Moore (o la falacia definista) excluye todas las definiciones, por ejemplo la de 'bueno', entonces excluye la definición de cualquier término. Para que sea efectivo de alguna manera se ha de interpretar como diciendo 'Cada término significa lo que significa y no lo que viene significado por otro término'. El señor Moore parece que implícitamente entiende su lema de esta manera en la Sección 13, pues procede como si 'bueno' no tuviera significado alguno, como si no tuviera significado único alguno. Si se toma el lema de esta manera, se seguirá que 'bueno' es un término indefinible, pues no se le pueden hallar sinónimos. Pero se seguirá también que no hay término que lo sea, y entonces el método de análisis es tan inútil como un carnicero inglés en un mundo sin ovejas.

Quizá hemos mal interpretado la falacia definista. Y ciertamente algunos de los pasajes que he citado anteriormente en este mismo artículo parecen suponer que la falacia naturalista es simplemente el error de definir una característica indefinible. Según esta interpretación, una vez más, la falacia definista en todas sus formas no tiene conexión especial con la bifurcación de lo ético y de lo no-ético. De nuevo, se puede cometer la falacia definista sin violar esa bifurcación, como cuando se define el placer en términos de rojo o la bondad en términos de correcto (si se concede la creencia del señor Moore de que el placer y la bondad son indefinibles). Pero incluso cuando se incurre en la bifurcación y se define la bondad en términos de deseo, el *error* no está en que se incurre en la bifurcación al violar (3), sino sólo en que se está definiendo una característica indefinible. Ello es posible porque la proposición de que la bondad es indefinible es independiente lógicamente de la

proposición sobre que la bondad es no-natural, como se muestra por el hecho de que una característica puede ser indefinible y con todo ser natural, como ocurre con lo amarillo, o no-natural y no obstante definible, como sucede con correcto (si se aceptan los puntos de vista del señor Moore acerca de lo amarillo y de lo correcto).

Consideremos la falacia definista tal cual la hemos planteado. Es sin duda un error definir una cualidad indefinible. Pero, de nuevo, la cuestión es ésta: ¿qué cualidades son indefinibles? Es una *petitio quaestionis* en favor del intuicionismo decir de antemano que la cualidad bondad es indefinible y que por tanto, todos los naturalistas cometen esa falacia. Se tiene que saber de antemano que la bondad es indefinible, si se quiere alegar que la falacia definista *es* una falacia. Entonces, sin embargo, la falacia definista puede entrar sólo al final de la controversia entre intuicionismo y definismo, y no se podrá usar como arma en la controversia.

La falacia definista se puede plantear de tal manera que abarque la bifurcación entre el 'debe' y el 'es' [29]. En tal caso, la cometería cualquiera que brindara alguna definición de cualquier característica ética en términos de características no-éticas. El inconveniente con tal definición, según esta interpretación, sería que se reduciría una característica *ética* a otra *no-ética*, y una *no-natural* a otra *natural*. Es decir, se excluiría la definición por el hecho de que la característica que se define es ética o no-natural y, por ende, no se puede definir en términos no-éticos o no-naturales. Pero, según esta interpretación, existe también el peligro de la *petitio* en la argumentación intuicionista. Suponer que la característica ética es exclusivamente ética es sin más pedir la cuestión de lo que está en tela de juicio cuando se brinda la definición. Así, de nuevo, se tiene que saber de antemano que la característica es no-natural e indefinible en térmi-

[29] Ver J. Wisdom, *Mind* (1931), p. 213, nota 1.

nos naturales, para poder afirmar que los definistas están cometiendo error.

El señor Moore, McTaggart y otros a veces formulan la falacia naturalista de manera algo diversa a las aquí tratadas. Dicen que los definistas confunden una proposición sintética universal acerca *del bien* con la definición de *bondad* [30]. El señor Abraham la llama 'falacia de una proposición mal construida' [31]. Aquí, de nuevo, la dificultad está en que, mientras es erróneo construir una proposición sintética universal como definición, para los intuicionistas es una *petitio* decir que aquello que los definistas están tomando como definición, en realidad es una proposición sintética universal [32].

Al final, empero, se esclarece cada vez más la situación entre intuicionistas y definistas (naturalistas o metafísicos). Todos los definistas sostienen que ciertas proposiciones que contienen términos éticos son analíticas, tautológicas o verdaderas por definición; *v. gr.*, el señor Perry considera así la proposición 'Todos los objetos de deseo son buenos'. Los intuicionistas sostienen que tales proposiciones son sintéticas. Lo que subyace en esta diferencia de opinión es que los intuicionistas proclaman tener al menos una oscura conciencia de una cualidad simple única o relación de la bondad o de lo correcto que aparece en la región que indican borrosamente nuestros términos éticos, mientras que los definistas alegan no poseer conciencia en absoluto de ninguna de esas cualidades y relaciones que pertenezcan al mismo contexto aunque se designen con palabras diferentes de 'bueno' y 'correcto' y sus sinónimos más obvios [33]. Los definistas afirman con toda sinceridad que sólo hallan una

30 Ver *Principia Ethica*, pp. 10, 16, 38; *The Nature of Existence*, vol. ii, p. 393.

31, Leo Abraham, 'The Logic of Intuitionism', *International Journal of Ethics*, vol. ii, p. 398.

32 Como señala el señor Abraham, *loc. cit.*

33 Ver R. B. Perry, *General Theory of Value*, p. 30; *cf. Journal of Philosophy* (1931), p. 520.

característica donde los intuicionistas dicen encontrar dos; como el señor Perry alega sólo encontrar la propiedad de ser deseado, donde el señor Moore ve ésta y la propiedad de ser bueno. Se trata, pues, de algo que hace referencia a la inspección o intuición y versa sobre la conciencia o discernimiento de cualidades y relaciones [34]. Por esto no es posible decidir la cuestión sirviéndose de la noción de falacia.

Si hemos de tomar la palabra de los definistas, entonces en realidad no están confundiendo dos características entre sí, ni definen una característica indefinible, ni confunden definiciones y proposiciones universales sintéticas; en breve, no están cometiendo la falacia naturalista o definista en ninguna de las interpretaciones arriba dadas; pues la única falacia que cometen —la verdadera falacia naturalista o definista— es el fracaso en columbrar las cualidades y relaciones que son centrales en moral. Pero esto no es ni falacia ni confusión lógicas. Ni es propiamente un error, sino más bien cierto tipo de ceguera, análoga a la ceguera para los colores. También se puede atribuir este tipo de ceguera moral a los definistas sólo si tienen razón en su afirmación de que no poseen conciencia de características éticas únicas, y si los intuicionistas tienen razón al alegar la existencia de tales características. Pero dar a esto el nombre de 'falacia', incluso en un sentido lato, no tiene ni propósito ni está bien.

Por otra parte, si no existen tales características en los objetos a los que adscribimos predicados éticos, entonces los intuicionistas, si podemos tomarles la palabra, adolecen de una alucinación moral correspondiente. Los definistas pueden tachar a esto de falacia intuicionista o moralista, pero tiene tan poco de 'falacia' como la ceguera de que acabamos de hablar. De todas formas, no creen en la insistencia de los intuicionistas respecto a que sólo ven características

34 Ver H. Osborne, *Foundations of the Philosophy of Value*, páginas 15, 19, 70.

éticas únicas y, consecuentemente, no les atribuyen esta alucinación. Por su parte, simplemente deniegan que los intuicionistas hallen tales cualidades o relaciones únicas, y buscan algún modo plausible de dar razón del hecho de que haya gente muy respetable y digna de confianza que crea verlas [35]. Así, acusan a los intuicionistas de verbalismo, hipostización y de cosas por el estilo. Pero esta parte del asunto no nos incumbe ahora.

Lo que nos debe ahora ocupar es el hecho de que los intuicionistas no creen en la afirmación de los definistas. Se verían muy desconcertados si realmente tuvieran que pensar que sus opositores tienen ceguera moral, pues no creen que sea preciso haber sido regenerados por la gracia para poder poseer discernimiento moral, sino que juzgan que la moralidad es algo democrático, aunque no todos los hombres sean buenos. Sostienen que si no 'todos advertimos' ciertas características únicas cuando empleamos los términos 'bueno', 'correcto', etc., es por falta de clareza analítica de la mente, inducida quizá por un prejuicio filosófico que no permite percatarnos en modo alguno de que son diferentes de otras características de las que sí nos percatamos [36]. Ahora bien, he estado sosteniendo que los intuicionistas no pueden tachar a los definistas de cometer falacia alguna, a menos que —y hasta que— demuestren que todos, incluidos los definistas, somos conscientes de las características únicas objeto de disensión. Si, a pesar de todo, lograran demostrar tal cosa, entonces y al final de la controversia podrían acusar a los definistas del error de confundir dos características, o del error de definir una característica indefinible, y estos dos errores podrían recibir el nombre de 'falacias', puesto que este vocablo es algo laxo en sus hábitos, aunque no se trataría de falacias lógicas en el sentido que lo es una argumentación no válida. La falacia de la pro-

35 *Cf.* R. B. Perry, *Journal of Philosophy* (1931), pp. 520 *ss.*
36 *Principia Ethica*, pp. 17, 38, 59, 61.

posición mal construida dependerá del error de confundir dos características y, por ende, en nuestra suposición presente, podría atribuirse también a los definistas, pero en realidad no se trata de una confusión *lógica*[37], puesto que no comporta confusión acerca de la diferencia entre proposición y definición.

Mas es difícil ver cómo pueden probar los intuicionistas que los definistas se percatan siquiera vagamente del requisito de las características únicas[38]. Esta cuestión se ha de dejar a la inspección o intuición de los definistas mismos, ayudados de las sugerencias que sean y que los intuicionistas consideren apropiadas. Así podremos dar crédito al veredicto de su inspección, especialmente al de aquéllos que hayan leído con ponderación los escritos de los intuicionistas, pero entonces de lo único que podrán ser acusados será de ceguera moral.

Además de intentar descubrir qué se entiende por falacia naturalista, me he esforzado en mostrar que la noción de que los definistas cometen una falacia lógica o cuasi-lógica no hace más que confundir las instancias entre intuicionistas y definistas (y las instancias entre estos últimos y los emotivistas o postulacionistas) y distorsiona el modo como debería plantearse la cuestión. En el procedimiento de los definistas no tiene por qué aparecer falacia alguna, ni siquiera se tiene que echar mano de falacias de sentido menos estricto para fallar el caso en contra de los definistas; a lo más, se podrán atribuir a los definistas sólo después de haber decidido el caso en su contra en campos independientes. Pero el único defecto atribuible a los definistas, *si* los intuicionistas tienen razón en afirmar la existencia de características éticas únicas indefinibles, es una ceguera moral peculiar, que no es falacia ni siquiera en sentido

37 Pero vèr H. Osborne, *op. cit.*, pp. 18 *s.*

38 Para una breve discusión de sus argumentos, ver *ibid.*, p. 67; L. Abraham, *op. cit.* Creo que todos son inconcluyentes, mas no lo puedo demostrar aquí.

lato. La cuestión debe decidirse mediante cualquier método que juzguemos satisfactorio para determinar si una palabra equivale o no a una característica y, si así es, si equivale a una característica única. Cuál sea el método a emplear es quizá, de una forma u otra, el problema básico de la filosofía contemporánea, pero no se ha llegado aún a alguna solución que en general sea satisfactoria. Sólo me atreveré a afirmar lo siguiente: me parece que no se ha de fallar en contra de los intuicionistas aplicando *ab extra* a los juicios éticos dictado alguno de significación empírica u ontológica [39].

39 Ver *Principia Ethica*, pp. 124 *s.*, 140.

IV

BIEN Y MAL[1]

P. T. Geach

De *Analysis*, vol. 17 (1956), pp. 33-42. Reimpreso con venia del autor, de *Analysis* y de Basil Blackwell.

Empezaré haciendo una distinción lógica entre dos clases de adjetivos, basada en la que existe entre adjetivos *atributivos* (*v. g.*, 'libro rojo') y adjetivos *predicativos* (*v. g.*, 'este libro es rojo'); tomo esta terminología de los gramáticos. Diré que en una frase 'BA' (donde 'B' es un nombre y 'A' un adjetivo), 'A' es un adjetivo (lógicamente) predicativo si la predicación 'es un BA' se escinde lógicamente en un par de predicaciones: 'es un B' y 'es un A'; en otro caso diré que 'A' es un adjetivo (lógicamente) atributivo. De aquí en adelante emplearé los términos 'adjetivo predicativo' y 'adjetivo atributivo' siempre en mi especial sentido lógico, a menos que se indique lo contrario con el adverbio 'gramaticalmente'.

Hay ejemplos de todos conocidos de lo que llamo adjetivos atributivos. 'Grande' y 'pequeño' son atributivos; 'x es una pulga grande' no se escinde en

1 [Este artículo lo discute A. Duncan-Jones, 'Good Things and Good Thieves'. *Analysis* (1966). Hacen al caso también P. R. Foot, 'Goodness and Choice', *Aristotelian Society Supplementary Volume*, XXXV (1961); Z. Vendler, 'The Grammar of Goodness', *Philosophical Review* (1963), y T. E. Patton y P. Ziff, 'On Vendler's Grammar of «Good»', *Philosophical Review* (1964). E.]

'x es una pulga' y 'x es grande'; ni 'x es un elefante pequeño' se escinde en 'x es un elefante' y 'x es pequeño', puesto que si tales análisis fueran legítimos, con un argumento simple se podría demostrar que una pulga grande es un animal grande y que un elefante pequeño es un animal pequeño. De nuevo, el tipo de adjetivo que los medioevales llamaban *alienans* es atributivo; 'x es un billete falso' no se escinde en 'x es un billete' y 'x es falso'; ni 'x es el padre putativo de y' se divide en 'x es el padre de y' y 'x es putativo'. Por otra parte, en la frase 'un libro rojo', 'rojo' es un adjetivo predicativo en mi sentido, aunque no es tal gramaticalmente, puesto que 'es un libro rojo' se escinde lógicamente en 'es un libro' y 'es rojo'.

Ahora puedo enunciar mi primera tesis sobre bueno y malo: 'bueno' y 'malo' son siempre atributivos, no predicativos. Esto se ve bastante claramente con 'malo', puesto que 'malo' es como un adjetivo *alienans*, pues no podemos predicar seguramente de un mal A lo que predicamos de A, lo mismo que no podemos predicar de un billete falso o de un padre putativo lo que predicamos de un billete o de un padre. Es a la moneda falsificada a la que llamamos 'mala', y no podemos inferir, *v. g.*, que pues el alimento da vida, el mal alimento también la da. No es tan claro el asunto, a primera vista, cuando se trata de 'bueno', puesto que 'bueno' es no *alienans:* todo lo que es cierto de un A lo es de un buen A. Pero consideremos el contraste en este par de frases 'coche rojo' y 'coche bueno'. Podría asegurar que un objeto lejano es un coche rojo porque puedo ver que es rojo, y un amigo de vista aguzada pero daltoniano podría ver que es un coche; pero no es posible asegurar que una cosa es un buen coche mancomunando noticias independientes sobre que es bueno y que es un coche. Este ejemplo nos muestra que 'bueno', lo mismo que 'malo', son esencialmente adjetivos atributivos. Incluso cuando 'bueno' y 'malo' hacen las veces de predicados, y por lo tanto gramaticalmente son predica-

tivos, se ha de entender algún sustantivo; no existe nada que sea bueno o malo, sino un buen o mal individuo. (Si digo que algo es una *cosa* buena o mala, o 'cosa' es un lugarteniente sin más de un nombre más apropiado que se ha de entender por el contexto, o estoy empleando 'bueno' o 'malo' como predicados y el que gramaticalmente sean atributivos es un simple disfraz. Según mi tesis, este último caso es ilegítimo.)

Ciertamente, se puede decir *simpliciter* 'A es bueno' o 'A es malo', si 'A' es nombre propio; pero se trata de una excepción que prueba la regla. Pues Locke estaba equivocado sin duda al afirmar que no existe esencia nominal de los individuos; el uso continuado de un nombre propio 'A' presupone siempre una referencia continuada a un individuo que es el mismo X, donde 'X' es un nombre común, y 'X' expresa la esencia nominal de un individuo llamado 'A'. Así, el empleo del nombre propio 'Peter Geach' presupone la referencia continuada al mismo *hombre;* el empleo de Támesis, la continuada referencia al mismo *río,* y así sucesivamente. En los libros de lógica moderna se lee con frecuencia que los nombres propios no tienen significado, en el sentido de 'significado' según el que se dice que lo tienen los nombres comunes; o (más oscuramente) que carecen de 'connotación'. Pero consideremos la diferencia entre la inteligencia que alguien tiene de una conversación que ha oído en una alquería cuando sabe que 'Lucio' es un hombre y cuando no sabe si 'Lucio' es un hombre, un arroyo o un perro. En el primer caso sabe *qué* significa 'Lucio', aunque no sepa *quién es;* en el segundo caso no sabe *qué* significa 'Lucio' y le es imposible seguir el curso de la conversación. Bien, pues, si el nombre común 'X' expresa la esencia nominal de un individuo llamado 'A', si *ser el mismo X* es condición cuyo cumplimiento viene presupuesto por el hecho de continuar llamando 'A' a un individuo, entonces el significado de 'A es bueno/malo', dicho *simpliciter,* será 'A es un buen/mal X'. Por ejemplo, si 'Lucio' es

un hombre, 'Lucio es bueno' dicho *simpliciter* significará 'Lucio es un buen hombre', aunque el contexto le dé el significado de que 'Lucio es un buen cazador' o cosa parecida.

Los filósofos morales conocidos como objetivistas [2] admitirán todo lo que he dicho correspondiente a los usos ordinarios de los términos 'bueno' y 'malo'; pero argumentarán que existe un uso predicativo de esos términos esencialmente distinto en expresiones como 'el placer es bueno' y 'preferir el gusto al deber es malo', y que éste es el único empleo de importancia filosófica. Para los objetivistas, los empleos ordinarios de 'bueno' y 'malo' no son más que una maraña de ambigüedades. Leí una vez un artículo, escrito por un objetivista, donde exponía estas ambigüedades y los desastrosos efectos que causaban en los filósofos cuando no las tenían en cuenta. Filósofo descarriado a este propósito fue Aristóteles. Aristóteles no hablaba inglés, pero por notable coincidencia ἀγαθός poseía ambigüedades paralelas a las de 'good' (bueno). Siempre pueden ocurrir tales coincidencias y hasta a veces se pueden traducir los equívocos. Pero también es posible que las acepciones de ἀγαθός y de 'bueno' sean paralelas, porque expresan idéntico concepto; concepto filosóficamente importante en el que Aristóteles hizo bien en interesarse, y porque la disclución aparente de este concepto en una masa de ambigüedades proviene de asimilarlo a los conceptos expresados por los adjetivos predicativos ordinarios. Es prejuicio sin más pensar que, o todas las cosas llamadas 'buenas' satisfacen una sola condición, o el término 'bueno' es irremisiblemente ambiguo. El filósofo que se deje en el tintero la mayoría de las acepciones de 'bueno' como trivialidades del inglés, puede resumirlas con algún acierto diciendo que expresan alguna condición definida que cumplen las cosas buenas; *v. g.*, que o contienen o que condu-

2 [Parece que Geach tenía en mente a Moore y a Ross; quiza también a Prichard. E.]

cen al placer; o bien, que satisfacen el deseo. Tales teorías de lo que es la bondad están expuestas a bien conocidas objeciones: son casos de falacia naturalista, como la llaman los objetivistas. La teoría de éstos es que 'bueno', en las acepciones escogidas que dejan a la palabra, no rinde una descripción ordinaria y 'natural' de las cosas, sino que les adjudica un atributo simple y *no*-natural, indefinible. Pero nadie ha dado jamás una razón coherente e inteligible de lo que ha de ser un atributo para que se le pueda considerar no-natural. Me temo que los objetivistas estén en un tira y afloja con el vocablo 'atributo'. Para poder asimilar 'bueno' a los adjetivos predicativos ordinarios, como 'rojo' y 'dulce', llaman atributo a bondad, y para zafarse de consecuencias indeseadas, provenientes de la asimilación, pueden protestar siempre: '¡Oh, no, no es eso! La bondad no es atributo *natural* como lo rojo y la dulzura, sino que es un atributo no-natural'. Es como si alguien quisiera esquivar la fuerza de los argumentos de Frege sobre que el número 7 no es una cifra, diciendo que lo es, pero sólo es una cifra no-natural, posibilidad que habría pasado por alto Frege.

Por otra parte, ¿puede un filósofo brindar expresiones filosóficas como 'el placer es bueno', para *explicar* cómo se ha de entender 'bueno' en sus discusiones? 'Deja a un lado las acepciones que «bueno» tiene en el lenguaje ordinario', dice el objetivista; 'en nuestra discusión querrá decir lo que quiero decir cuando empleo expresiones cual «el placer es bueno». Ya se imagina, es claro, cómo quiero que éstas se tomen, pero no, no puedo explicarme más; ¿no ve que «bueno», en mi sentido, es un término simple e indefinible?' Pero, ¿cómo se nos puede pedir que desde el principio demos por sentado que una acepción filosófica peculiar tenga que significar necesariamente algo en absoluto? Menos se ha de esperar que, ya desde el principio, sepamos qué significa esa acepción.

Concluyo que el objetivismo es sólo un intento de librarse de la falacia naturalista; en realidad no nos

explica cómo 'bueno' difiere en su lógica de otros vocablos, sino que sólo ofusca con palabras que no conocemos.

Lo que hasta aquí he dicho tendrá reconocimiento general entre los escritores de ética contemporáneos de Oxford (a los que en adelante denominaré los moralistas oxonienses). Consideraré ahora su razonamiento positivo sobre 'bueno'. Sostienen que las características del vocablo que hasta aquí he tratado derivan de que su función es primordialmente comendatoria y en modo alguno descriptiva. 'Este es un buen libro' significa más o menos 'recomiendo este libro' o 'escoge este libro'. Defienden, sin embargo, que si bien la fuerza primordial de 'bueno' es la comendatoria, hay muchos casos en que su fuerza es puramente descriptiva; 'Hutton batió un buen tanto', en un reportaje periodístico no significaría: '¡Qué maravilloso tanto batió Hutton! ¡Quién lograra un tanto así cuando fuera su turno!' Los moralistas oxonienses explican tales casos diciendo que 'bueno' está, por así decir, entre comillas: Hutton batió un 'buen' tanto; es decir, un tanto que los buenos jugadores de cricket llamarían 'bueno', o sea, que comendarían y escogerían.

Rechazo por completo la opinión de que 'bueno' no tiene primordialmente sentido descriptivo. Alguien a quien le importara un comino el cricket, pero que entendiera perfectamente las reglas del juego (suposición no imposible), daría a la frase 'batir un buen tanto' sentido puramente descriptivo, independientemente de los gustos de los aficionados al cricket. Así, cuando digo que alguien es un buen atracador o un buen asesino, no lo estoy comendando. Es posible imaginar situaciones en que estas descripciones podrían servir para guiar la elección que otro tuviera que hacer (por ejemplo, cuando el jefe de un comando debiera escoger asaltantes y asesinos para un cometido especial); pero estas circunstancias son raras

3 [Se ha alterado ligeramente el texto, para salvar la mala inteligencia que surgió en la primera versión. E.]

y no dan el sentido primario de las descripciones. Debe quedar en claro que llamar a una cosa una buena A no influye en la elección, a menos que quien haya de escoger quiera una A; pero esta influencia sobre la acción no es la fuerza lógicamente primaria de la palabra 'bueno'. 'Tienes hormigas en los pantalones', que sin duda primariamente tiene fuerza descriptiva, afectará más de cerca la acción que otras muchas acepciones del término 'bueno'. Y hay muchas acepciones de la palabra 'bueno' que no hacen referencia a los gustos de una mesa de expertos o de algo parecido. Si digo que un hombre tiene buen ojo o buen estómago, mi indicación claramente es descriptiva y no hace referencia a ninguna mesa de especialistas en ojos o estómagos.

Por lo que puedo colegir de sus escritos, los moralistas oxonienses presentarían dos líneas de objeciones contra la opinión de que 'bueno' primariamente tiene fuerza descriptiva. En primer lugar, si salvamos los errores gemelos de la falacia naturalista y del objetivismo, veremos que no existe descripción alguna, 'natural' o 'no-natural', a que correspondan todas las cosas buenas. Los rasgos por los que se llama 'buena' a una cosa difieren según sea la cosa en cuestión. Se dice que un cuchillo es 'bueno' si tiene UVW; se dice que un estómago es 'bueno' si posee XYZ, etc. Así, pues, si 'bueno' tuviera una fuerza propiamente descriptiva, ésta variaría de caso a caso: 'bueno' aplicado a cuchillos expresaría los atributos UVW; 'bueno' aplicado a estómagos expresaría los atributos XYZ, y así sucesivamente. Si 'bueno' no ha de ser ambiguo sin más, se ha de suponer que su fuerza primordial será invariablemente comendatoria, mas no la indefinidamente variable fuerza descriptiva.

Esta argumentación es una falacia sin más. Se trata de otro ejemplo de asimilación de 'bueno' a adjetivos predicativos ordinarios o, más bien, supone que esta asimilación estaría perfecta si la fuerza de 'bueno' fuera descriptiva. No se seguiría de hecho, aun en el caso de que 'bueno' fuera un adjetivo predicativo

ordinario, que si 'buen cuchillo' significara lo mismo que 'cuchillo que es UVW', 'bueno' significaría lo mismo que 'UVW'. Triángulo de lados iguales significa lo mismo que 'triángulo con tres lados iguales', pero no se puede borrar 'triángulo' y decir que 'con tres lados iguales' significa lo mismo que 'de lados iguales'. En el caso de 'bueno', la falacia es todavía más burda; es como pensar que 'cuadrado de' significa lo mismo que 'doble de', porque 'el cuadrado de 2' es lo mismo que 'el doble de 2'. Esta analogía matemática puede ayudarnos a aclarar ideas. No existe número alguno por el que se pueda multiplicar otro para que dé su cuadrado; pero de aquí no se sigue ni que 'cuadrado de' sea una expresión ambigua que a veces significa 'doble de', 'triple de', etc., ni que se deba hacer algo diferente que multiplicar para hallar el cuadrado de un número; pero, dado un número, ya tenemos su cuadrado. De manera similar, no existe descripción alguna a la que respondan todas las cosas llamadas 'buen tal'; pero no se seguirá tampoco ni que 'bueno' sea expresión muy ambigua o que llamar buena a una cosa sea algo que difiera de su descripción; y dada la fuerza descriptiva de A, la fuerza descriptiva de 'un buen A' no dependerá de los gustos de la gente.

'Pero podría saber qué significaría «buen higrómetro», aunque no supiera para qué sirven los higrómetros. En tal caso, sin embargo, no podría dar fuerza descriptiva definida a «buen higrómetro», como algo contrapuesto a «higrómetro»; por tanto, «bueno» ha de poseer fuerza recomendatoria y no descriptiva'. La respuesta a esta objeción (imitada de argumentos reales de los moralistas oxonienses) es que si no sé para qué sirven los higrómetros, no sé qué significa 'higrómetro'; yo sólo sé que podría averiguar su significado indagando para qué sirven, de igual manera como sé cómo podría hallar el cuadrado de los habitantes de Sark, si supiera el número de ellos, y *en tal sentido* podría decir que entiendo la frase 'el cuadrado del número de habitantes de Sark'.

La segunda línea de objeciones de los moralistas oxonienses consiste en preguntar primero si la conexión entre llamar a una cosa 'una buena A' y aconsejar a alguien que necesite una A que elija ésta, es analítica o empírica, y, luego, establecer un dilema. Parece del todo equivocado pensar que la conexión es un hecho meramente empírico; pero si decimos que es analítico, entonces 'bueno' no puede tener fuerza descriptiva, pues no se puede inferir lógicamente de una mera descripción.

Debería, pues, decir que la conexión no es meramente empírica, pero tampoco es analítica. Pertenece a la *ratio* de 'querer', 'escoger', 'bueno' y 'malo' que, normalmente, y siendo iguales otras cosas, alguien que quiera A, escogerá una buena A y no una mala A; o más bien, que escogerá una A que piensa ser buena y no una A que crea que es mala. Esto vale tanto si las Aes que escogemos son cuchillos como caballos o ladrones; *quidquid appetitur, appetitur sub specie boni*. Puesto que la frase cualificante 'normalmente, y siendo iguales otras cosas' es necesaria para la verdad de esta proposición, no es analítica. Pero la presencia de estas frases no reduce la proposición a una mera generalización empírica burda; pensar esto sería cometer una cruda falacia empirista que reveló de una vez por todas Wittgenstein. Incluso si no todas las Aes son Bes, la proposición de que las Aes normalmente son Bes puede pertenecer a la *ratio* de A. La mayoría de las tiradas de ajedrez son válidas, la mayoría de las intenciones se ejecutan, la mayoría de las proposiciones son verídicas... ninguna de estas proposiciones es sin más una generalización burda, pues si tratáramos de describir qué ocurriría si la mayoría de las tiradas de ajedrez fueran inválidas, que la mayoría de las intenciones no se ejecutaran y que la mayoría de las proposiciones fueran mentiras, resultaría que no haríamos sino hablar sandeces. Hablaríamos sandeces si tratáramos de describir a un pueblo cuya costumbre fuera escoger las Aes malas

cuando desearan Aes, rechazando las Aes buenas. (Y esto vale para todas las interpretaciones de A.)

Admito que se encuentra mayor dificultad en pasar de 'hombre' a 'buen/mal/hombre' o de 'acto humano' a 'buen/mal/acto humano', si estas frases se han de tomar como puramente descriptivas y en sentidos determinados simplemente por los de 'hombre' y 'acto humano'. Creo que podría salvarse esta dificultad, pero aun así los moralistas oxonienses podrían desplegar ahora un argumento que sería poderosa arma. Supongamos que hemos hallado un significado claramente descriptivo de 'buen acto humano' y de 'mal acto humano' y que hemos demostrado que el adulterio responde a la descripción 'mal acto humano'. ¿Por qué esta consideración habría de disuadir a un adúltero volente? ¿Por qué paso lógico podríamos ir del sentido supuestamente descriptivo 'el adulterio es un acto humano malo', al imperativo 'no cometerás adulterio'? Es inútil decir 'Es deber tuyo hacer el bien y evitar el mal'; esto o es idéntico a la observación inútil 'Es bueno hacer el bien y evitar hacer el mal', o 'Es deber tuyo' es intromisión de una fuerza imperativa que no conllevan los términos 'bien' y 'mal' los cuales, *ex hypothesi*, son puramente descriptivos. En primer lugar tenemos que conceder que la pregunta '¿Por qué he de?' o '¿Por qué no he de?' es racional y merece respuesta y no observaciones destempladas sobre la maldad de interrogar, y creo que la única respuesta pertinente es algo que apele a lo que el preguntante *quiere*. Desde la época de Kant, la gente ha supuesto que existe otro tipo de respuesta pertinente: que apele no a la inclinación, sino al Sentido del Deber. Ahora bien, se podría lograr que alguien, mediante educación, recabara un estado de la mente según el cual 'No *has de*' fuera respuesta suficiente a '¿Por qué no he de?'; en el cual estado, al darse esta respuesta a sí mismo (o al escuchar cómo se la dan otros), quede embargado de un temor harto peculiar, y en el que hasta esté convencido de que 'no debe' preguntar siquiera por qué 'no debe'. (Cf. el

poema juvenil de Lewis Carroll 'My Fairy' con su devastadora 'Moral: No debes'.) Los filósofos morales de la escuela objecionista —como el señor David Ross— llamarían a esto 'aprehensión de las propias obligaciones'; no les preocupa que, bendito sea Dios, este tipo de educación pueda hacer que un individuo llegue a 'aprehender' prácticamente todo como sus 'obligaciones'. (En efecto, admiran al hombre que hace lo que piensa *debe* hacer, independientemente de lo que realmente haga: ¿no está actuando por el Sentido del Deber, que es el motivo más excelso?) Pero, aun dado el caso que 'no debes' sea respuesta definitiva, *ad hominem*, a '¿Por qué no he de?', no es en absoluto respuesta racional.

Creo que se me puede demostrar que el hecho de que una acción sea una acción humana buena o mala es de por sí algo que toca los deseos del agente. Aunque llamar a una cosa 'una buena A' o 'una mala A' no actúa por sí sobre los deseos del oyente, puede esperarse que sí lo haga si el oyente está escogiendo una A. Ahora, lo que un hombre no puede dejar de escoger es su manera de actuar; así, llamar a una manera de actuar buena o mala sólo puede servir para guiar la acción. Como dice Aristóteles, el actuar bien, εὐπραξία, es la meta del hombre *simpliciter*, ἁπλῶς y *qua* hombre. Hay otros objetos de elección que lo son sólo relativamente, πρός τι, o son objetos de un hombre particular, τινός[4]; pero todo hombre ha de escoger cómo actuar, por lo que llamar a una acción buena o mala no dependerá, para su efecto suasorio, de las peculiaridades individuales de deseo.

No me dedicaré aquí a explicitar el vigor descriptivo de 'acción humana buena (mala)', pero parece conveniente hacer ciertas observaciones sobre la lógica de la frase. En primer lugar, un raquetazo del tenis o una tirada de ajedrez es un acto humano; ¿hemos de decir, pues, que la descripción 'buen raquetazo' o 'buena tirada' es de por sí algo que apela

4 E. N. 1139*b* 2-4.

a los deseos del agente? Es claro que no; pero no hay dificultad aquí. Aunque el raquetazo y la tirada sean actos humanos, no se sigue que un buen raquetazo o una buena tirada sean buenos actos humanos por la lógica peculiar del término 'bueno'. Así, llamar a un raquetazo o a una tirada buenos no es *eo ipso* apelar a lo que un agente pueda desear.

En segundo lugar, aunque podamos hablar con sentido de un buen o mal acto humano, no podemos hablar con sentido de un buen o mal acontecimiento o de una buena o mala cosa que ha sucedido. 'Acontecimiento', lo mismo que 'cosa', son palabras demasiado vacías para comportar ya un criterio de identidad ya un estándar de bondad. Preguntar: '¿Es bueno o malo (que suceda) esto?' es tan inútil como preguntar: '¿Es esto lo mismo que vi ayer?' o '¿Es el mismo acontecimiento?', a menos que se llene la vacuidad de 'cosa' o 'acontecimiento' mediante algún contexto de expresión. El asesinato de César fue una cosa mala para un organismo viviente, un buen sino para un hombre que deseaba la latría, y —de manera similar— un buen o mal acto por parte de sus asesinos; mas preguntar si fue un acontecimiento bueno o malo carecería de sentido.

En tercer lugar, estoy pasando por alto deliberadamente la supuesta distinción entre lo Correcto y lo Bueno. En santo Tomás de Aquino no hay tal distinción. Le basta en hablar de actos humanos buenos y malos. Cuando Ross diría que hay una acción moralmente buena, pero no un acto correcto, el Aquinate diría que una buena intención humana había resultado en lo que fue, de hecho, una mala acción; y cuando Ross diría que había una acción correcta, pero no una acción moralmente buena, Tomás de Aquino diría que había un mal acto humano efectuado en circunstancias en que un acto similar con intención diferente habría sido bueno (*v. gr.*, dar dinero a un menesteroso por vanagloria y no para socorrerle).

Puesto que la palabra inglesa 'right' (correcto) tiende por el genio del idioma a tomar el artículo definido

—hablamos de *una* (*a*) buena tirada, pero de *la* (*the*) tirada correcta—, quienes crean que actuar correctamente es algo diferente de actuar bien, considerarán que la conducta virtuosa consiste, no precisamente en actuar bien y evitar actuar mal, sino en hacer en cada ocasión *el* acto correcto en esa circunstancia. Esta doctrina especialmente estricta conduce de hecho a consecuencias muy laxas. Quien procure actuar bien y evitar actuar mal, si sabe que el adulterio es un acto malo, concluirá que (como dice Aristóteles) no será posible deliberar cuándo o cómo o con quién cometer adulterio [5]. Pero quien crea que en cada ocasión hay que discernir *el* acto correcto en esa circunstancia, puede llegar a la conclusión de que en esta ocación, consideradas bien todas las cosas, el adulterio es *la* acción correcta. Sir David Ross nos dice explícitamente que puede darse que *el* acto correcto sea la condenación judicial del inocente 'para que no perezca toda la nación', puesto que en ese caso 'el deber *prima facie* de consultar el interés general resulta más obligatorio que el deber *prima facie* claramente perceptible de respetar los derechos de aquéllos que han respetado los derechos ajenos' [6]. (Hemos de esperar caritativamente que las palabras de Caifás, que cita, tienen sólo el halo reverente de un texto bíblico, y que no tuvo presente qué asesinato judicial se estaba aconsejando.) [7]

Sé muy bien que mucho de esta discusión no satisface. No he podido desenvolver con la debida extensión algunos puntos donde creo ver claro; hay otros puntos (por ejemplo, la relación entre el deseo y el bien, y la *ratio* precisa del mal en los actos ma-

5 E. N. 1107*a* 16.

6 *The Right and the Good*, p. 61.

7 Algunos de los que sostienen la noción *del* acto correcto, piensan incluso que *el* acto correcto de un Dios debería ser creativo; v. g., que un Dios estaría obligado a crear el mejor de los mundos posibles, de modo que este mundo nuestro o es el mejor de los posibles o no existe Dios bueno. No me adentraré en esto; bastará decir que lo que se ha de esperar de un buen Creador es *un* buen mundo; no *el* mundo correcto.

los) donde no veo claro. Además, si he alegado que la característica de que sea buena o mala una acción humana influirá de por sí los deseos del agente, no he tratado sobre si una acción mala por naturaleza se ha de evitar siempre y en toda circunstancia, como creía Aristóteles. Pero quizá, si no he dejado claras todas las cosas, habré dejado más claras algunas.

V

GEACH: BIEN Y MAL

R. M. Hare

De *Analysis*, vol. 13 (1957), pp. 103-12. Reimpreso con la venia del autor y de *Analysis*. Aparte algunas pequeñas alteraciones a los párrafos segundo y tercero, correspondientes a cambios en el artículo del profesor Geach, la respuesta del profesor Hare se ha reimpreso sin revisión.

El señor Geach me ha sugerido que publique una réplica a su artículo 'Bien y Mal' [1]. Concluyo de esto que me considera parte constituyente de esa compleja Tía Sally [*] que denomina 'moralistas oxonienses'. Con todo, no me toca defender ese monstruo heterogéneo. En la batalla escénica que Geach sostiene con su criatura me veo en ambos bandos, pues si bien algunos de los puntos de los 'moralistas oxfordianos' son versiones más o menos reconocibles de los míos, también lo son bastantes de los propios argumentos, y de algunos ejemplos, de Geach. No voy a atacar

1 *Analysis*, vol. 17, núm. 2, pp. 33-42. Deseo agradecer al señor Geach la amabilidad que tuvo al proporcionarme el manuscrito a máquina de un artículo más extenso, del cual el que aquí aparece es la sección introductoria; así como el aclararme, por escrito, el significado que da a *ratio*, y el empleo que quiere dar a ese concepto en su teoría.

* Monigote con aspecto de mujer y con una pipa que en las ferias inglesas sirve de blanco; la suerte está en hacerle caer la pipa. (T.)

su tesis principal de que 'bueno' es un adjetivo atributivo, puesto que concuerdo con él [2].

Cuán compleja criatura sea la Tía Sally de Geach se puede ver considerando un párrafo típico de su escrito, el tercero —completo— de la p. 36 [3]. Se dice allí que los moralistas oxfordianos mantienen las siguientes posiciones:

(1) La función de 'bueno' primariamente no es descriptiva en absoluto, sino comendatoria.
(2) 'Este es un buen libro' significa algo como 'Te recomiendo este libro'.
(3) 'Este es un buen libro' significa algo como 'Escoge este libro'.

Puede ser que Geach no haya advertido la diferencia entre comendar y recomendar, [4] o entre estos términos y los distintos fines con los que se emplea el imperativo, o entre estos variados particulares y las dos cosas diferentes que se expresan con las proposiciones 'Qué tanto maravilloso estaba batiendo Hutton' u 'Ojalá pudiera siempre uno batir semejantes tantos', que se las compone para embutir en su revesada bolsa. Puesto que su último ejemplo está sacado de *LM*, p. 118, diríase que, según ese libro, el significado comendatorio de 'bueno' se ha de identificar con una expresión exclamativa o con el deseo. Pero ese punto de vista no aparece en el texto del libro. Si Geach

2 Esta tesis ha sido común entre los moralistas oxonienses durante muchos años. Por lo que recuerdo, la oí por primera vez cuando se discutía a Frege con el profesor Austin. En *Foundations of Arithmetic* (a cargo y traducido por el profesor Austin, pp. 28 *ss.*), Frege, siguiendo una sugerencia de Baumann, señala que los números cardinales son atributivos, en el sentido de Geach. Hay que conceder reconocimientos a H. W. B. Joseph y, en definitiva, a Aristóteles, *Eth. Nic.*, I, 6. Esta tesis se encuentra, aunque sin esta terminología, en mi *Language of Morals (LM)*, p. 133.

3 [Párrafo 4, p. 97 de este volumen. Ed.]

4 Según el *O.E.D.* [*Oxford Etymological Dict.*], 'comendar' se emplea a veces con el sentido de 'recomendar'; pero este uso no es común, ni es en este sentido como la palabra aparece en *LM*. Empleamos normalmente 'recomendar' cuando se trata de una *elección* particular, y 'comendar' cuando se menciona algo en general como 'digno de aceptación o aprobación'.

desea atribuir estas confusiones a otros, además de atribuírselas a sí mismo, ¿no debería citarlos?[5]

Tampoco veo claro por qué se ha de pensar que los 'moralistas oxonienses', al hallarse frente al ejemplo del 'buen tanto (wicket)', habrían de emplear el argumento que Geach pone en sus bocas. Geach presenta el ejemplo como un caso en que 'la fuerza de «bien» es puramente descriptiva'. Los 'moralistas oxonienses', dice Geach, 'explican tales casos diciendo que aquí «bueno» se ha de decir entre comillas; Hutton estaba batiendo un «buen» tanto, es decir, un tanto que los aficionados al cricket llamarían «bueno», o sea, que comendarían elegir'. Hay casos por cierto en que 'bueno' se emplea con comas invertidas (' ')[6], pero no éste tal. Se trata más bien de casos en que la palabra 'bueno' no tiene significado valuatorio, porque el hablante no está comendando, sino que sólo alude a la comendación de otra gente (de ordinario bien conocida). Pero en este caso, el escritor está comendando sin duda el tanto (aunque no está haciendo otras cosas que Geach confunde con comendar). En este contexto, cabalmente, el propósito primero al decir en un informe periodístico que se trató de un buen tanto es 'informar a los lectores qué tipo de tanto era'[7]. Pero se puede suponer con seguridad que el cronista y la mayoría de sus lectores son aficionados al cricket y que, por tanto, aceptan la norma de comendación que va adherida a la frase. Si por el mismo uso común la frase no conllevara esa norma o estándar de comendación,

5 No reclamo lugar alguno en estas posiciones atribuidas a los 'Moralistas oxfordianos'. A mi manera de ver, 'bueno' normalmente tiene tanto el sentido descriptivo como el valuatorio (comendatorio), siendo primario el valuatorio. Hay que distinguir esta posición de 1) arriba, donde las palabras 'en absoluto' *parecen* suponer que la palabra no tiene 'primariamente' (signifique esto lo que signifique) ninguna acepción en absoluto, sino comendación; esta última posición la rechazo específicamente en *LM*, pp. 121 *s.*

6 Ver *LM*, p. 124.

7 Ver *LM*, p. 118.

no podría emplearse para informar. Además, el razonamiento de Geach se basa en la presunción de que se puede demostrar que el sentido de una expresión no es primordialmente valuatorio si se trae *un* contexto en el que se emplee con finalidad primariamente descriptiva. Difícil es hallar argumento más débil. Es raro también que Geach haya llegado a pensar que alguien que entendiera bien el juego pudiera 'dar un sentido puramente descriptivo a la frase «tanto bien batido» sin tomar en cuenta los gustos de los aficionados al cricket'. ¿Quiere decir que los estándares según los cuales se aplica esta frase a los tantos (wickets) no tiene nada que ver con las preferencias de los batidores?

Otro ejemplo de la confusión existente en la mente de los 'moralistas oxonienses' puede hallarse en el pasaje que inmediatamente sigue. Es claro que no distinguen entre decir que llamar a una cosa una buena A es *guiar* la elección, y decir que es *influir* o *afectar* la elección. La comendación puede proponerse guiar la elección, pero ciertamente no es de necesidad que se pretenda influir o afectar la elección [8]. No es (como diría Geach) parte de la *ratio* de la palabra 'bueno', o de la palabra 'comendar', o incluso de los imperativos, que las proposiciones 'bueno' (las comendaciones o los imperativos) tengan influencia causal sobre nuestra conducta. El ejemplo geachano de 'hormigas en los pantalones' es una objeción contra tal teoría, pero no es concluyente de

8 En mis artículos 'Imperative Sentences', *Mind* (1949), esp. 39, he tratado de aclarar esta distinción, lo mismo que en 'Freedom of the Will', *Ar. Soc. Supp.* Vol. XXV, esp. pp. 205-15,y en *LM*, pp. 13-15. El doctor Falk hace distinciones similares en 'Goading and Guiding', *Mind* (1953), p. 145; lo mismo el profesor Cross, 'The Emotive Theory of Ethics', *Ar. Soc. Supp.*, Vol. XXII, esp. pp. 139 y *ss.* Pero Cross no trata el asunto muy completamente, y me parece que Falk coloca imperativos de la línea divisoria donde no debiera. Quizá la cosa se esclarezca más si el profesor Austin publica, y cuando ello ocurra, algo sobre su distinción entre la fuerza ilocucionar y perlocucionar (es decir, entre lo que hacemos *al decir P*, y lo que intentamos hacer *diciendo* P).

por sí. En realidad se trata de una versión vulgarizada de un ejemplo que yo empleé en el primero de los artículos citados para mostrar esto en el caso de los imperativos: 'Si quieres que alguien se quite los pantalones, lo lograrás más fácilmente diciéndole «se te está subiendo un alacrán por la manga del pantalón», que diciéndole «Bájate los pantalones»'. Algunos filósofos, como el distinguido moralista de Cambridge y de Ann Arbor, profesor Stevenson, han sostenido que tanto los juicios morales como los imperativos son *de ratione* afectantes a la acción; otros, como el doctor Falk, han defendido que los imperativos sí lo son, no así los juicios morales. Ciertamente es cuestionable decir que lo son los juicios morales, y en esto estoy acorde con Geach. Pero, de nuevo, si cree que este punto de vista cuestionable es corriente en Oxford, ¿no debería citar a sus fautores?

En breve, para ser descriptivista (que es quizá el mejor calificativo de lo que soy) no es preciso ser emotivista en modo alguno; y en particular no es necesario ser un emotivista que confunde los juicios morales con la propaganda. Quizá, si Geach reflexiona sobre esta distinción, 'recomendar' no le causará en el futuro mayor intranquilidad que 'bueno'. Puesto que, una vez que se ha aclarado la mala inteligencia, desaparece también la razón principal para dudar de qué dice el *O. E. D.* * sobre 'bueno'. La primera cosa que ese diccionario dice sobre el significado de 'bueno' es que se trata 'del adjetivo más general de recomendación'. El hecho de que Sir David Ross (a quien nadie puede superar como adicto descriptivista) copie esta definición sin disentir, fortalece el ligamen entre 'bueno' y recomendar[9]. Y, en efecto, se vuelve realmente muy difícil negar esta asociación si consideramos lo que el mismo diccionario dice sobre la palabra 'recomendar'. La define como 'mencionar como digno de aceptación o de aprobación'; 'aprobar' se define

* *Oxford Etymological Dictionary.*

9 Ver *The Right and the Good*, p. 66.

como 'pronunciar como bueno, recomendar'. Si conjuntamos esas dos definiciones, tenemos: 'Recomendar: mencionar como digno de...ser pronunciado como bueno', o brevemente, 'mencionar como bueno'. Si esto es lo que significa 'recomendar', ¿cómo va a ser tan impropio, como Geach evidentemente piensa, decir que 'bueno' tiene como función primaria recomendar?

En este punto se puede objetar que, aunque el diccionario tenga plena razón en conectar 'bueno' con 'recomendar', como lo hace, me equivoco en dar el paso ulterior de unir recomendar con la guía de las elecciones. Tal objeción podría poner cualquiera que a toda costa quisiera mantener 'bueno' como palabra puramente descriptiva, a pesar de su conexión (que difícilmente puede negarse) con recomendar. Pero este argumento no va con Geach, pues en las pp. 38 *s.* de su artículo [10] dice 'Pertenece a la *ratio* de «querer», «escoger», «bueno» y «malo», el que normalmente y siendo iguales las demás cosas, quien desee una A escogerá una A que crea buena y no elegirá una A que juzgue mala'. No hay duda de que Geach tiene razón al decir que la doctrina del *quidquid appetitur, appetitur sub specie boni*, no es analítica tal cual aparece, 'puesto que la frase cualificativa «normalmente y siendo iguales las demás cosas» se requiere necesariamente para que la proposición sea cierta'. Pero si a la proposición se le añade esta frase cualificativa, se convierte no meramente en verdadera sino también en analítica, y esto es todo cuanto se requiere para demostrar que el significado de la palabra 'bueno' no es *puramente* descriptivo.

Mi principal propósito en este artículo, al que ahora me vuelvo, es estimar la sugerencia de Geach sobre hasta qué punto 'bueno' posee fuerza descriptiva. Que *tiene* fuerza descriptiva lo he dicho muchas veces, pero Geach quiere ir más allá. Mientras que yo sos-

10 [p. 107 de este volumen. Ed.]

tengo que el significado común a todos los ejemplos del empleo de la palabra no puede ser descriptivo y que su significado común hay que ir a buscarlo en la función valuatoria (comendatoria) de la palabra, Geach sostiene que este significado común es una especie de significado descriptivo. Así, piensa, 'bueno' tiene el mismo significado descriptivo en las expresiones 'buen cuchillo' y 'buen estómago', aunque —como los dos coincidimos— 'las características por las que se llama «buena» a una cosa son diferentes según el tipo de cosa en cuestión [11]. Piensa que esto puede ser así porque, aunque no existan características comunes, el significado de la palabra 'bueno', tomado conjuntamente con el de la palabra 'cuchillo' o con el de la palabra 'estómago', nos permite especificar los rasgos que han de tener las cosas de este tipo para que se las pueda llamar 'buenas'. Compara esto con el modo como, aunque no tengamos que multiplicar 2 por el mismo factor para tener su cuadrado, cual se debe hacer con 3 para sacar *su* cuadrado, no obstante la expresión 'el cuadrado de' tiene un significado común; dado un número, queda determinado su cuadrado [12].

Me di perfecta cuenta de esta línea de argumentación cuando escribí *LM*, pp. 99-103, y en ese pasaje se contienen las consideraciones que a mi modo de ver dan respuesta al caso. Existe cierta clase de palabras (llamadas 'palabras funcionales' en *LM*) con las que cabe seguir ese mismo procedimiento. 'Una palabra es funcional si, para explicar plenamente su significado, hemos de decir *para qué* es el objeto a que se refiere, qué se supone que ha de hacer [13]. Son ejemplos de palabras funcionales 'berbiquí', 'cuchillo' e 'higrómetro'. Las definiciones que de estas palabras

11 p. 37 [p. 106 de este volumen. Ed.]

12 Este ejemplo da pie a reflexión muy útil. En la p. 36 de *LM* se encontrará material para dicha reflexión, pues ocurre allí ejemplo similar. Para comparar la conexión entre mi empleo del ejemplo y el de Geach, véase abajo, p. 108, n. 1 [p. 80. Ed.]

13 *LM*, p. 100; cf. Geach, p. 38. [p. 107. Ed.]

traen los diccionarios hacen referencia a las funciones de los objetos así llamados. Por tanto, si sabemos el significado de 'bueno' y también el de 'higrómetro', estaremos en disposición de saber qué características ha de poseer un higrómetro para que se le pueda llamar bueno (de hecho, sabemos muy bien cuál es uno de los rasgos que nos autorizarían a llamarlo malo, a saber, que soliera registrar un contenido de humedad de un gas diferente del que realmente posee).

Cuando 'bueno' se refiere a una palabra funcional, es cierto la mayor parte de lo que dice Geach. Sin embargo, pasa de manera acrítica de esta verdad sobre las palabras funcionales a un aserto más arrollador (lo que está injustificado), a saber, que vale decir lo mismo de todos las acepciones de 'bueno'. Esto es lo que debería mostrar, si desea asentar su pretensión de que el significado común de 'bueno' es descriptivo. 'Bueno' con frecuencia se refiere a palabras que no son funcionales. En tales circunstancias, para poder saber qué características ha de tener la cosa en cuestión para que se la pueda llamar buena no basta con saber el significado de la palabra. Necesitamos saber, además, qué norma hemos de adoptar para juzgar la bondad de esa clase de cosas, y la norma no se nos manifiesta, siquiera en parte (como con las palabras funcionales), con el significado de la palabra que va con 'bueno'. Así, tenemos que saber no sólo el significado de 'bueno', sino también el significado de 'puesta del sol' (y también saber el significado de toda la expresión 'buena puesta del sol'), sin que por ello se nos den los rasgos que ha de tener la puesta del sol para que se la pueda llamar buena. Hay, es cierto, alguna concordancia, entre los que gustan contemplar la puesta del sol, sobre qué puestas pueden llamarse buenas (tienen que ser brillantes, sin deslumbrar, y cubrir buena parte del firmamento de colores variados e intensos, etc.), pero esta norma no va indicada en el significado de 'puesta' y mucho menos en el de 'buena'.

Se ha de recalcar que esta diferencia entre el comportamiento de 'bueno' cuando se refiere a un sustantivo funcional y su conducta cuando se refiere a un nombre no funcional no se debe a diferencia alguna en el significado de 'bueno' en sí. Podemos decir, más o menos, que en los dos casos significa 'tener las cualidades características (sean cuales sean) comendables en el tipo de objeto en cuestión'. La diferencia en los dos casos está en que la palabra funcional nos da las pistas sobre cuáles son esas cualidades; no así la palabra no funcional. Ello es porque al clasificar algo como un higrómetro, *v. g.*, hemos determinado ya que su valoración estribará en que se amolde a cierto estándar; no así cuando clasificamos algo distinto, como la puesta de sol. Así, la palabra 'higrómetro', a diferencia de la palabra 'puesta', no es puramente descriptiva. Para saber el significado de 'higrómetro', no sólo debemos saber qué propiedades observables ha de poseer algo para que se le pueda denominar higrómetro, sino que debemos saber también qué nos justificará que comendemos o condenemos algo como higrómetro. Nada de esto vale en el caso de 'puesta'. Para conocer el significado de 'puesta' nos basta con saber que podemos dar ese nombre a lo que vemos en el firmamento de poniente cuando, a todas vistas, el sol se hunde en el horizonte [14].

Es obvio que la intención de Geach era que cuanto dice de 'bueno' en general podía ser aplicable a las acepciones morales de la palabra. Surge la cuestión de si las palabras que acompañan a 'bueno', en contextos morales, son siempre palabras funcionales. A mi modo de ver, la mera concurrencia de una palabra funcional junto con 'bueno' normalmente es indicativo de que el contexto *no* es moral. Existen algunas excepciones posibles a esta regla; así, *v. g.*, la frase

[14] La explicación de la paradoja de que la expresión 'buen higrómetro' tiene un significado *descriptivo* fijo, precisamente porque las dos palabras que lo componen son a la vez *valuatorias*, en parte se hará evidente a todo el que compare *LM* pp. 100-101 con *ibid.* pp. 36-7; las dos valoraciones se 'anulan recíprocamente'.

'buen ejemplo', ocurre en contextos morales, y 'ejemplo' en tales contextos es posiblemente una palabra funcional que significa 'cosas dignas de imitarse'. No sé por seguro qué razón se ha de dar de tal expresión, pero afortunadamente en esta argumentación no necesito sostener que en contextos morales 'bueno' nunca se emplea con palabras funcionales, sino sólo que a veces se emplea con palabras no-funcionales. Pues habré mostrado con ello que, en todo caso, en esos contextos ni 'bueno' en sí, ni la expresión total en que ocurra son puramente descriptivos. Y así habré demostrado que, si existe un significado común de 'bueno' que se dé en todos los casos, la razón que de este común significado nos ofrece Geach es inadecuada.

'Es un buen hombre' es un juicio moral en algunos contextos, aunque no en otros. Si 'hombre' se emplea (como ocurre a veces) significando 'soldado' o 'criado' (ambas palabras funcionales), la expresión 'buen hombre' es no-moral, precisamente porque la palabra 'hombre' se emplea de manera funcional. Es parte de las definiciones de soldado o de criado, que posean ciertos deberes; el criado que actúe contra los deseos o intereses de su patrón es *eo ipso* un mal sirviente, y el soldado cuya conducta lleva a la pérdida de la guerra de los suyos es *eo ipso* un mal soldado. Pero si 'hombre' se emplea de la manera ordinaria y general para indicar 'miembro de la especie humana' no es funcional, y es así como se emplea en los contextos morales. Creo que lo mismo vale para la expresión 'buena acción humana' que Geach emplea; pero puesto que esta expresión no es de uso común, es difícil estar seguro. De todas formas, en la expresión común 'buena acción', 'acción' no es funcional. Se puede saber el significado de 'acción' sin saber nada que determine, siquiera en grado mínimo, qué acciones se han de calificar como 'buenas' o 'malas'. Y si 'humano', al igual que 'hombre', no es una palabra funcional, lo mismo valdrá para 'acción humana'.

No obstante, no es necesario para mi argumento hacer suposiciones sobre qué se incluye o deja de incluirse en el significado de la palabra 'hombre'. Bastará con considerar distintas cosas que podrían ir incluidas, y advertir las consecuencias lógicas de su inclusión. Como ocurre con frecuencia en filosofía, no hay nada aquí que dependa del uso corriente y *real* de los vocablos, pero si acordamos emplearlos de una manera determinada, deberemos atenernos a las consecuencias. Podemos acordar entender por 'hombre' simplemente 'criatura viviente de la siguiente forma física...' especificando esa forma. Si esto fuera lo que entendiéramos por 'hombre', claramente la palabra no sería funcional y toda la expresión 'buen hombre' tampoco sería descriptiva. Pero estaría dispuesto a concordar con Geach si protestara que por 'hombre' entendemos más que eso. Pues, como me ha indicado, podría haber criaturas que tuvieran la misma forma del hombre, pero a las que no podríamos adscribirles ese nombre porque carecieran de ciertas capacidades intelectuales, como el habla racional. Es cierto que llamamos 'hombre' a un vástago de padres humanos que carece de esa atribución, pero si descubriéramos una *raza* de criaturas que carecieran de esa potestad, podríamos vacilar en si les otorgaríamos ese nombre.

Hasta aquí, podemos concordar Geach y yo; pero una cosa es decir que al llamar a una criatura hombre suponemos que pertenece a una especie que posee ciertas *capacidades*, y otra decir que *bueno específico* [15] es de cierto tipo. Por ejemplo, podríamos rehusarnos a atribuir el nombre de 'hombres' a una especie de criaturas que, si por lo demás fueran iguales que los hombres que conocemos, psicológicamente fueran incapaces de mentir, de matar o de cometer ninguna de las demás cosas que vulgarmente se llaman pecaminosas. Podríamos decir 'no son humanos, sino que más bien deberíamos llamarles «ángeles»,

15 Tomo esta expresión de una carta de Geach.

o (en caso de que se nos objetara por parte de la teología) con un nuevo nombre que las señalara distintamente'. Si fuera según eso como empleáramos la palabra 'hombre', la posesión de esos *poderes* (de mentir, asesinar, etc.) sería parte de la *ratio* de la palabra 'hombre' así empleada. Pero de esto no se seguiría que el ejercicio de esos poderes, o incluso su posesión, fuera conductora hacia el bien específico del hombre, o que desbaratarlos o refrenar su ejercicio (por ejemplo, mediante una educación moral total) fuera contrario al bien específico del hombre [16].

Si Geach desea hacer posible la derivación de conclusiones del significado de 'hombre' sobre qué es contrario o proficuo a que un hombre sea buen hombre, debería incluir en el significado de la palabra no sólo ciertas estipulaciones acerca de las *capacidades* de quienes aspiran a la designación de 'hombre', sino también algo sobre lo que es ser *buen* hombre. En breve, deberá convertir 'hombre' en vocablo funcional. Supongamos ahora que Geach se toma esa libertad. Entonces, toda la expresión 'buen hombre' y quizá expresiones tales como 'buena acción humana' recibirán significados descriptivos fijos. Pero habrá pagado serio gravamen por esa franquicia. Significaría que lo dicho en su p. 40 [17] ya no sería cierto: 'Lo que un hombre no puede dejar de escoger es su manera de actuar; así, llamar a una manera de actuar buena o mala sólo puede servir para guiar la acción'.

16 Geach es el último de una famosa sucesión de pensadores que han confundido sistemáticamente 'lo que una cosa *puede* (o, alternativamente, *puede típicamente* o *hace típicamente)* hacer', con la noción, del todo diferente, 'lo que una cosa *debe* hacer (o, alternativamente, lo que es específicamente *bueno* para ella)'. Platón fue el culpable principal. Se ha empleado la palabra 'función' para abarcar todas estas nociones, pero sólo se justifica su asimilación si aceptamos la precisa *Natura (sive Deus) nihil facit inane* [La Naturaleza (o Dios) nada hace en vano (T)]. Todo aquél que se sienta atraído por el empleo que Geach hace de este tipo de raciocinio debería leer primero a Aristóteles, *Política*, 1252 a 35, donde se emplea una premisa similar para justificar la esclavitud y la sumisión de las mujeres (cf. también p. de este volumen. Ed.)

17 [p. 97 de este volumen. Ed.]

En la definición sugerida de 'hombre', y por ende de 'humano', ya no ocurriría esto, si 'acción' (como supone Geach en la primera línea del párrafo de donde proviene esta cita) es un modo abreviado de decir 'acción humana'. Pues al elegir qué debo hacer puedo estar escogiendo, no dentro de la clase de comparación 'acciones humanas', sino dentro de otra clase más vasta. Similarmente, si 'caballo' se emplea como palabra funcional, como 'corcel' [caballo de guerra], el caballo que eche al suelo a su jinete es *eo ipso* un mal caballo, pero el *caballo* podría decirse 'No tengo intención de ser caballo en *ese sentido;* yo soy sólo un cuadrúpedo solípedo perisodáctilo (Equus caballus), con largas crines y cola', y dedicarse a echar al jinete, sin pecar contra nada, salvo contra los estándares de éste. Pues si el significado de la palabra 'corcel' determina alguna de las cualidades de un buen corcel, el de la palabra 'caballo', según la definición más general dada por el *O. E. D.*, no las determina; en este sentido de 'caballo' queda abierta la cuestión de cómo han de comportarse los caballos a su modo de ver. Precisamente *porque* el caballo no puede hacer menos que ser caballo en este sentido general, el hecho de que sea un caballo en este sentido general no determina si debería escoger ser un buen corcel, o dejar de hacerlo. Podría considerar la elección que se le ofrece no como una elección sobre qué tipo de corcel debe ser, sino sólo —más generalmente— qué tipo de caballo ha de ser. Fácil sería el oficio de picador si los caballos se pudieran convertir en corceles por definición.

VI
CREENCIAS MORALES [1]

Philippa Foot

De *Proceedings of the Aristotelian Society*, Vol. 59 (1958-9), páginas 83-104. Reimpreso por cortesía de la autora y del editor de Aristotelian Society.

Para muchos, el adelanto más notable en filosofía moral de los últimos cincuenta años ha sido la refutación del naturalismo, y se sorprenden si en fechas tan tardías se vuelve a abrir la causa. Es fácil entender su actitud; dadas ciertas suposiciones incuestionables, resulta tan poco sensato hablar del naturalismo como proponerse la cuadratura del círculo. Quienes lo ven de esta manera se contentan con el sobreaviso de que toda teoría naturalista ha de poseer una trampa y sólo les inquieta tener que desperdiciar más tiempo buscando la falacia antigua. Este artículo pretende persuadirles de que reconsideren críticamente las premisas sobre las que se basan sus argumentos.

No sería exagerado decir que toda la filosofía moral, cual se enseña hoy por doquier, estriba en un contraste entre proposiciones de hecho y valoraciones; algo que más o menos suena así: 'La verdad o falsedad de las proposiciones de hecho queda patente mediante las pruebas, y lo que sirve de prueba

1 [Este artículo ha sido criticado, *v. g.*, por M. Tanner, 'Examples in Moral Philosophy', *Proceedings of the Aristotelian Society* (1964-5); D. Z. Phillips, 'Does it Pay to be Good?', *ibid.*; D. Z. Phillips, 'On Morality's Having a Point', *Philosophy* (1965). Ed.]

está en el significado de las expresiones que aparecen en la proposición de hecho. (Por ejemplo, el significado de «redondo» y «plano» hizo que las pruebas aportadas por los viajes de Magallanes fueran en favor de la redondez y no de la horizontalidad de la Tierra; quienquiera que hubiese seguido cuestionando si las pruebas eran pruebas podría haber sido convencido de que estaba cometiendo un error lingüístico.) Se sigue que dos personas no pueden proferir la misma proposición y traer como pruebas cosas distintas; al cabo, uno de ellos al menos podría ser convencido de ignorancia lingüística. Se sigue también que si a un hombre se le presentan buenas pruebas respecto de una conclusión fáctica, no puede rechazar sin más la conclusión basado en que en su esquema de cosas esas pruebas no son pruebas en absoluto. Con las valoraciones el asunto está de otro modo. La valoración no está conectada lógicamente con proposiciones de hecho sobre las que se base. Una persona puede decir que una cosa es buena por algún hecho que ella posea, mientras que otra puede rehusarse a tomar ese hecho como prueba, pues no hay nada contenido en el significado de «bueno» que lo conecte con una «prueba» y no con otra. Se sigue que un excéntrico en moral puede objetar las conclusiones morales a partir de premisas muy idiosincrásicas; podría decir, *v. g.*, que un hombre es buen hombre porque apretara o separara las manos y nunca se encaminara hacia el NNE. después de haberse dirigido al SSO. Hasta podría rechazar la valoración que al respecto hiciera otro, negando simplemente que sus pruebas lo fueran en modo alguno.

'El hecho acerca de «bueno», que hace que el excéntrico pueda servirse del término sin caer en el tremedal de la falta de significado, es su función «guía de la acción» o «práctica». Es algo inalienable, pues como cualquier otro se cree en la disponibilidad de escoger las cosas que llamn «buenas» y dejar las que llame «malas». Como los demás, emplea «bueno» sólo en conexión con una «pro-actitud»; lo que ocurre

es que él posee pro-actitudes para cosas muy distintas, y por eso las llama buenas.'

Existen respecto de la 'valoración' dos presuposiciones, que denominaré presuposición (1) y presuposición (2).

La presuposición (1) es que, sin error lógico, un individuo podría basar sus creencias respecto a asuntos de valor, sobre premisas en que los demás no verían pruebas de nada. La presuposición (2) es que, dado el tipo de proposición que los demás consideran como probatoria respecto de una conclusión de valor, un individuo puede rehusarse a sacar la conclusión porque *ello* no cuente como prueba para *él*.

Consideremos la presuposición (1). Podemos decir que dependerá de la posibilidad de mantener sólido el significado de 'bueno' a través de todos los cambios en los hechos referentes a cualquier cosa que haya de contar en favor de su bondad. (No quiero indicar que alguien pueda hacer cambios con la misma rapidez con que escoge, sino sólo que haya escogido lo que haya escogido, no es posible desmentirle.) Pero existe una formulación mejor que ataja disputas triviales sobre el significado que 'bueno' pueda tener el algún sector de la comunidad. Digamos que la presuposición consiste en que la función valuatoria de 'bueno' puede permanecer constante a través de los cambios que ocurran en el principio valorativo. Sobre esta base, podría decirse que incluso si nadie puede llamar *bueno* a un hombre porque junte o abra las manos, puede comendarlo o expresar su *pro-actitud* hacia él y, si es preciso, inventar un nuevo vocabulario moral que exprese su insólito código moral.

Aquellos que mantengan esta teoría añadirán a ella distintas cualificaciones, como es natural. En primer lugar, la mayoría está de acuerdo con Hare, y en contra de Stevenson, sobre que tales palabras como 'bueno' sólo se adjudican a casos individuales mediante la aplicación de principios generales, de modo que incluso el excéntrico moral más extremado ha de

aceptar principios de comendación. En segundo lugar, 'comendar', 'tener una pro-actitud', etc., se supone que están relacionados con hacer y elegir, de manera que sería imposible decir, *v. g.*, que un hombre sería bueno sólo si viviera mil años. La escala de valoración es de suponer que ha de restringirse a la gama de acción y elección posibles. No es mi cometido poner en tela de juicio estas restricciones supuestas sobre el empleo de los vocablos valorativos, sino sólo contender que no son suficientes.

La cuestión crucial es ésta. ¿Es posible extraer del significado de palabras como 'bueno' algún elemento llamado 'significado valorativo' al que podamos considerar como externamente relacionado con sus objetos? Habría tal elemento, por ejemplo, si se diera la regla de que cuando se 'comendara' determinada acción, el hablante se sintiera obligado a aceptar el imperativo 'tengo que hacer estas cosas'. Tal cosa estaría relacionada externamente con su objeto, porque, dentro de la limitación que antes hemos advertido referente a las acciones posibles, tendría sentido pensar que existe algo que es sujeto de tal 'comendación'. Según esta hipótesis, un excéntrico moral podría comendar el apretar las manos como acción propia de hombre bueno, y no tendríamos que mirar ningún trasfondo para otorgar sentido a la presuposición. O sea, en esta hipótesis, el juntar las manos se podría comendar sin explicación alguna; podría ser lo que llaman quienes sostienen estas teorías 'principio moral último'.

He de dejar en claro que esta hipótesis es insostenible y que no es posible describir el significado valoratorio de 'bueno', valorar o comendar o nada por el estilo, sin fijar el objeto al que se supone que van adheridos. Sin poner antes las manos sobre el objeto propio de cosas como la valoración, sólo prenderemos en nuestra red, o algo muy diferente, como aceptar una orden o tomar una resolución, o nada en absoluto.

Antes de pasar a considerar esta cuestión, trataré otras actitudes mentales y creencias que poseen esta relación interna con su objeto. Con esto espero aclarar el concepto de relación interna con un objeto y, de paso, en caso de que mis ejemplos levanten resistencia pero sean por fin aceptados, mostrar cuán fácil es preterir una relación interna cuando la hay.

Veamos, por ejemplo, el orgullo.

La gente se suele asombrar si se le dice que existen límites respecto de las cosas de las que alguien puede sentirse orgulloso, de aquello que realmente pueda enorgullecerse. No sé a ciencia cierta qué entiendan por orgullo, si quizá algo que se refiere a sonreír y andar con aire garboso, o mantener en alto un objeto para que la gente lo pueda ver, o quizá piensen que el orgullo es un tipo de sensación interna, de manera que uno se sienta con deseos de golpearse el pecho diciendo 'el orgullo es algo que siento *aquí*'. Las dificultades que presenta este segundo punto de vista son bien conocidas; el objeto lógicamente privado no puede ser algún nombre que sea nombre de algo en el lenguaje público [2]. La primera manera de ver la cosa es la más plausible y puede parecer razonable decir que, dada cierta conducta, cabe describir a un hombre demostrando que está orgulloso de algo, que puede ser cualquier cosa. En un sentido esto es verdadero, aunque no en otro. Dada una descripción de un objeto, de una acción, de una característica personal, etc., no es posible descartarlo como objeto de orgullo. Antes de hacerlo, es preciso saber qué diría de ello un hombre que se sienta orgulloso o esté orgulloso de eso mismo; pero si no mantiene las creencias correctas acerca de ese objeto, entonces sea cual sea su actitud, no se trata de orgullo. Considérese, por ejemplo, la idea de que alguien pudiera estar orgulloso del firmamento o del mar: los contempla y lo que siente es orgullo, o hincha su pecho y gesticula

2 Ver Wittgenstein, *Philosophical Investigations*, especialmente §§243-315.

con orgullo dirigiéndose hacia ellos. Tendrá sentido esto sólo si se efectúa una presuposición especial respecto a sus creencias; *v. g.*, que es presa de una demencia y cree que ha librado al firmamento de su desmoronamiento o al mar de secarse. El objeto característico del orgullo es algo visto (*a*) como propio de alguna manera, y (*b*) como algún tipo de logro o ventaja; sin esto es imposible hablar de orgullo. Para convencerse de que el segundo punto es indispensable, se puede suponer que alguien se siente orgulloso porque ha colocado una mano sobre la otra tres veces en una hora. Otra vez, aquí el presupuesto de que es orgullo lo que siente tendrá buen sentido si se completa cierto trasfondo. Quizá está enfermo y es toda una hazaña hacer ese movimiento; quizá su ademán tiene importancia religiosa o política, o quizá es un hombre denodado y está desafiando a los dioses o a sus gobernantes. Pero si no existe trasfondo alguno, no puede haber orgullo, no porque nadie pudiera psicológicamente sentir orgullo en tal caso, sino porque fuera lo que fuera lo que sintiese no sería orgullo lógicamente. Sin duda, la gente ve cosas y empresas insólitas, aunque no cualquier cosa sin más, y se puede identificar con antepasados remotos, con deudos y con vecinos y hasta con la humanidad. No voy a negar que hay ejemplos de orgullo estrambóticos y cómicos.

Podríamos haber escogido otros muchos ejemplos de actitudes mentales que internamente se relacionan con su objeto de manera similar. Así, el miedo no es temblar, correr y volverse pálido; sin el pensamiento de que amenaza un daño, nada de todo eso constituirá miedo. Ni se podría decir que alguien ha sentido consternación ante algo que no ha visto como malo; si sus pensamientos al respecto fueron de que se trataba de algo del todo bueno, no podría afirmar que (cosa rara) lo que sintió fue congoja. '¡Qué raro, me sentí desanimado cuando debería haberme sentido contento!' es el preludio de la búsqueda del aspecto adverso de la cosa, que se cree acecha tras la fachada

placentera. Alguien objetará que el orgullo, el miedo y la angustia son sentimientos o emociones y que, por tanto, no se pueden tomar como analogía apropiada de la 'comendación', y que podría ser provechoso discutir otro tipo de ejemplo. Así, podríamos indagar la creencia de que cierta cosa es peligrosa y preguntar si cabría lógicamente mantener esto respecto de cualquier otra cosa. A la par que 'esto es bueno', la expresión 'esto es peligroso' es un aserto que naturalmente podríamos aceptar o rechazar comprobando su verdad o su falsedad. Parece que solemos apoyar tales asertos con pruebas y, además, diríase que existe una 'función monitora' relacionada con la palabra 'peligroso', como se supone que la hay 'comendatoria' con la palabra 'bueno'. Pues supongamos que los filósofos, confusos acerca de la propiedad de la peligrosidad, concluyeran que la palabra no significara propiedad alguna, sino que esencialmente fuera un término práctico o guía de la acción empleado para *avisar*. A menos que se emplee en un sentido 'entrecomillado', el vocablo 'peligroso' servirá para advertir; significando esto que quienquiera que así lo aplicara indicaría que evitaría las cosas que llamara peligrosas, o prevendría a los demás para que no se aproximaran a ellas, sino que corrieran en dirección opuesta. Si la conclusión no fuera obviamente ridícula, sería fácil inferir que quien aplicara el término diferentemente a como lo hacemos nosotros, podría decir que las cosas más peregrinas eran peligrosas, sin temor al rechazo; la idea sería que cabría afirmar que 'las consideraba peligrosas', o al menos como 'conminatorias', porque debido a su actitud y acciones habrían cumplido las condiciones para ser tales cosas. Esto es absurdo, porque, sin su debido objeto, el *aviso* —como *el creer en la peligrosidad*— no tendría razón de ser. Es lógicamente imposible advertir sobre algo que no se piensa como amenazante y malo, y para que haya peligro se requiere un tipo serio de perjuicio, como lastimarse o la muerte.

Existen, con todo, algunas diferencias entre pensar que una cosa es peligrosa y el sentirse orgulloso, atemorizado o entristecido. Cuando alguien dice que algo es peligroso ha de apoyar su afirmación con algún testimonio especial; pero cuando afirma que se siente orgulloso, atemorizado o desazonado, la descripción que haga del objeto de su orgullo, de su miedo o de su angustia, no es preciso que posea tanta relación con su afirmación primordial. Si se le demuestra que la cosa de que se siente orgulloso no es suya después de todo, o no fue asunto de tanta magnitud, tendrá que decir que su orgullo era injustificado, pero no ha de retirar la afirmación que hiciera de que estaba orgulloso. Por otra parte, alguien que diga que una cosa es peligrosa y posteriormente vea que cometió un error en pensar que de ella podría resultar perjuicio, debe volver sobre su aserto y confesar que estaba equivocado. Pero en ninguno de los dos casos, el hablante puede seguir como antes. Quien descubriera que no fue su calabaza sino la de otro la ganadora del premio sólo podría decir que se sentía todavía orgulloso, si pudiera presentar alguna otra razón de orgullo. Es de esta manera cómo incluso los sentimientos son vulnerables por los hechos.

Se objetará probablemente contra estos ejemplos que, por parte del modo al menos, hay *petitio quaestionis*. Se dirá que, en efecto, alguien sólo puede estar orgulloso de algo que considera como buena acción, como una proeza o un emblema de cuna noble, de la misma manera que sólo puede sentir desasosiego ante algo que ve como malo, y temor ante algún daño que le amenaza; similarmente, sólo podrá prevenir si puede hablar, pongamos por caso, de lesiones. Pero esto limitará el campo de objetos posibles de esas actitudes y creencias sólo si el campo de estos vocablos se limita a su vez. Para disipar esta objeción trataré el significado de 'lesión', puesto que es el caso más simple. El que se sienta inclinado a decir que todo podría considerarse una proeza, o como algo malo que atemorizara a la gente, o por

lo que ésta se sintiera descorazonada, debería probar lo siguiente. Quiero considerar la proposición de que cualquier cosa se podría tener por peligrosa, porque si causa lesiones es peligrosa, y cualquier cosa podría considerarse perjudicial. Pensaré en el daño corporal, porque es el tipo de perjuicio que se asocia con el peligro. No es razonable colocar un letrero junto a la carretera donde diga '¡Peligro!', porque haya matorrales que puedan rayar el coche, ni se puede rotular un producto como 'peligroso' porque pueda dañar tejidos delicados. Aunque podemos hablar del peligro de que así ocurra, no es ésta la acepción de la palabra que aquí considero.

Cuando un cuerpo se lesiona, se altera de manera especial, empeorando. Ahora nos interesa saber cuáles son las alteraciones que se consideran lesiones; antes que nada hay que tener en cuenta, por ejemplo, cómo sobrevienen las lesiones que no se deben a menoscabo natural. Parece claro que no cualquier cosa dejará, *v. g.*, marca insólita en el cuerpo que no se oblitere, por empeño que se tenga en hacerla desaparecer. El tipo de lesión más importante con mucho es la que atañe a alguna parte del cuerpo y obstruye su funcionamiento: lesiones en la pierna, en el ojo, en el oído, en la mano, en un músculo, en el corazón, en el cerebro o en la medula espinal. Si la lesión afecta el ojo, probablemente sufrirá detrimento la visión; si está en la mano, impedirá que se pueda extender y haga o realice otras funciones; la pierna puede lesionarse y quedar incapacitada para realizar sus movimientos y sostener la carga del cuerpo, y el pulmón puede debilitarse a tal grado que no logre inhalar la debida cantidad de aire. Cuando se trata de que no puede ser ejecutada una función correspondiente a una parte del cuerpo, no dudamos en hablar de lesiones, como en estos casos. Pero podemos dudar en decir que el cráneo puede recibir lesiones y a lo mejor preferimos hablar de daños que le pueden sobrevenir, debido a que, si bien el cráneo tiene una función, que es la protectora, carece de

operación. Pero al hablar de la función protectora del cráneo, a lo mejor ya no tenemos inconveniente en usar la palabra lesión. Si hacemos que el concepto de *lesión* dependa del de *función* queda aquél delimitado estrechamente, puesto que no cualquier uso al que se someta una parte del cuerpo contará como función. ¿A qué se debe que incluso cuando se trata de medios por los que algunos se ganan la vida, como la joroba del enano o la barba en la mujer, su remoción no la consideramos como lesión? Diremos que se trata de deformidades, pero no es por esto. Supongamos que hubiera un hombre con un músculo de más, imperceptible, en las orejas, que se ganara la vida, como bufón de corte, meneándolas; las orejas no sufrirían lesión si se hiciera desaparecer ese músculo. Si fuera natural la comunicación por medio de la oreja, entonces ésta tendría la función de señalar (no tenemos palabra para indicar este tipo de 'habla'), por lo cual un detrimento de esa función sería una lesión. Pero las cosas no son así. Ese bufón se serviría de las orejas para entretener a la gente, pero esa no sería la función de ellas.

No dudo de que mucha gente se impacientará porque se mencionen estos hechos, pues piensa ella que no tiene la menor importancia que *sucedan* cosas parecidas y le parecerá indiferente que la pérdida de la barba o de la joroba o de ese supuesto músculo de la oreja sea llamada o no lesión. Pero ¿no es catastrófico perder algo con lo que uno se gana la vida? Con todo, parece natural que tales particularidades no se cuenten como lesiones, si se toma en cuenta la condición de la vida humana y se contrapone la pérdida de una capacidad especial para hacer reír a la gente o lograr que se quede con la boca abierta, a la posibilidad de ver, oír, andar o asir las cosas. Lo primero se requiere para un modo de vivir muy especial, lo segundo sirve para todo. Esta distinción parece tanto más natural si consideramos qué otras amenazas, además de una lesión, pueden constituir peligro de muerte, por ejemplo, o de per-

turbación mental. Se llamará peligroso a un trauma que pueda ocasionar desequilibrio mental o amnesia, porque se necesita la inteligencia, la memoria y la concentración, lo mismo que la vista, el oído o el uso de las manos. No hablamos aquí de lesión, a menos que sea posible relacionar el detrimento con alguna alteración física; pero hablamos de peligro, porque existe la misma pérdida de una capacidad que todo hombre necesita.

Pueden existir lesiones fuera del ámbito que hemos estado ponderando, pues se puede decir que alguien ha recibido lesiones donde no se ha obstruido ninguna función somática. En general, creo que se puede hablar de lesión aunque se trate de un golpe que ocasionara dolor perdurable, sin que se sintiera ningún otro perjuicio, pero no conozco otra aplicación importante del concepto.

Parece, pues, que como la gama de cosas que pueden recibir el nombre de lesión es bastante limitada, la palabra 'peligroso' está también limitada por lo que se refiere a lo que causa lesión. Podemos afirmar que no se puede llamar peligrosa a cualquier cosa, por vallas que se erijan o por mucho que se gesticule.

Hasta aquí he estado sosteniendo que cosas tales como el orgullo, el temor, el descorazonamiento y el pensamiento de que algo puede ser peligroso poseen relación interna con su objeto, y espero que las cosas se vayan volviendo claras. Ahora debemos pasar a pensar si las actitudes o creencias que son incumbencia del filósofo moral son semejantes a las anteriores, si cosas como la 'valoración' o el 'considerar bueno algo' y la 'comendación' se podrían encontrar lógicamente en combinación con cualquier objeto. Todo lo que aquí puedo hacer es dar un ejemplo que haga inaceptable esa suposición y desbaratar los pocos soportes que pueda poseer. Pongamos como ejemplo las acciones triviales e insustanciales del hombre que apretara sus manos tres veces por hora, y digamos que a esto se le pudiera llamar buena acción. Nos abstendremos de añadir trasfondo alguno especial,

pues se ha de dejar bien claro que se trata de la cuestión sobre qué puede contar como bueno o malo en la acción de alguien, y no sobre qué se ha de pensar que es bueno o malo con un trasfondo especial. Creo que el enfoque de que hablo a veces parece plausible porque subrepticiamente se insinúa el trasfondo.

Quien dijera que el apretar las manos tres veces por hora era una buena acción debería responder antes a la cuestión '¿En qué sentido?', puesto que la proposición 'es una buena acción' no tiene significado claro. Hay que tener presente que, pues nuestro tema es la filosofía moral, aquí no significa 'fue una buena cosa', como se diría de algo que una persona hubiese llevado a cabo sensatamente en el curso de una gestión cualquiera, sino que debemos concentrar nuestra atención en el 'uso moral de «bueno»'. No veo claramente si tiene sentido hablar del 'uso moral de «bueno»', pero podemos citar un número de casos que suscitan cuestiones morales. Es porque éstas son tan variadas y porque 'ésta es una buena acción' no alude a ninguna de ellas, por lo que debemos preguntar '¿en qué sentido?'. Por ejemplo, hay cosas que si se ejecutan, cumplen un deber, como el deber de los padres para con los hijos o de éstos para con aquéllos. Supongo que cuando estos filósofos hablan de buenas acciones, incluyen también a éstas; otras pertenecen al capítulo de la caridad y se han de incluir también; otras son acciones que apelan a las virtudes del valor o de la templanza, y aquí el aspecto moral está en que se llevan a cabo a pesar del temor o de la tentación de placer. Se han de efectuar a causa de algún bien real o imaginario, pero no necesariamente por lo que los filósofos llamarían bien moral. El valor no se refiere *de manera particular* a poner en salvo las vidas ajenas, ni la templanza a permitir que los demás compartan alimento y bebida, sino que la bondad de lo *que se practica* puede ser todo tipo de utilidad. Es debido a que existen estos tan diversos casos incluibles (supongo) dentro de la expresión

'buena acción' por lo que nos abstenemos de aplicarla sin antes preguntar qué se quiere dar a entender, y así ahora deberíamos preguntar qué se pretende cuando alguien, supuestamente, dice que 'apretar las manos tres veces en el curso de una hora es una buena acción'. ¿Se supone que tal acción cumple un deber? Entonces, ¿en virtud de qué existe tal deber?, ¿o a quién se debe? Hemos prometido no insertar ningún trasfondo especial, pero posiblemente no tendrá el *deber* de apretar las manos, a menos que exista ese trasfondo. Ni puede ser un acto de caridad, pues no se ve que pueda beneficiar a nadie, ni tampoco un ademán de humildad, a menos que una presuposición especial la convierta en eso. La acción podría ser valerosa, pero sólo si se hiciera ante el miedo o por hacer el bien, pero no hay razón para inmiscuir circunstancias especiales que determinaran esto.

Estoy seguro de que no se planteará la siguiente objeción. 'Naturalmente, apretar las manos tres veces en el término de una hora no puede incluirse dentro de alguna de las virtudes que reconocemos como bien, pero esto sólo equivale a decir que no se trata de una buena acción según nuestro código moral corriente. Es lógicamente posible que, en un código moral distinto, se sancionaran virtudes por completo diversas, de las que no tenemos siquiera el nombre.' No puedo responder debidamente a esta objeción, pues para ello se debería dar una explicación satisfactoria del concepto de virtud. Pero cualquiera que piense que sería fácil describir una nueva virtud relacionada con el apretar las manos tres veces en una hora, debería intentarlo. Creo que se dará cuenta de que debe hacer trampa y suponer que en la comunidad respectiva se ha dado importancia especial al apretar las manos o se piensa que tiene algún efecto particular. La dificultad, obviamente, está en que sin trasfondo especial no existe posibilidad de responder a la cuestión '¿De qué se trata?'. No vale decir que habría razón en llevar a cabo la acción porque moralmente fuera una buena acción; el quid está en cómo

dar tal razón si antes no podemos hablar de qué se trata. Y es insensato suponer sin más que podamos imaginar que *cualquier* cosa es el asunto de que se trata, sin tener que aclarar *cuál* es el quid. Al apretar las manos se puede hacer un ligero sonido de succión, pero, ¿qué propósito tendría? Es sin duda claro que las virtudes morales tienen que estar conectadas con el bien o con el perjuicio humanos y que es imposible llamar bueno o perjuidicial a lo que uno quiera. Consideremos, por ejemplo, la suposición de que un hombre pudiera decir que había recibido daño porque se había sacado un cubo de agua del mar. Como siempre, se podrían hallar contingencias en que a esa expresión se le pudiera dar sentido; por ejemplo, si se relacionara con creencias mágicas, pero entonces el perjuicio estribaría en lo que hicieran los malos espíritus, no en el sacar agua del mar. Sería algo tan absurdo como si alguien dijera que se le ha perjudicado, porque le hubieran reducido los pelos de la cabeza a un número par [3].

Concluyo que la presuposición (1) es ciertamente muy dudosa y que no se puede hablar como si se entendiera qué es la 'valoración', la 'comendación' o las 'pro-actitudes' independientemente de las acciones en cuestión.

II

Deseo hablar ahora de lo que he llamado presuposición (2), según la cual alguien podría rehusar siempre asentir a la conclusión de un argumento so-

3 Ante esta clase de ejemplo, muchos filósofos se agazapan entre la maleza de la estética. Sería interesante saber si están dispuestos a dejar que todo su caso descanse en la posibilidad de que existieran objeciones de tipo estético a lo que se hiciera.

bre valores, porque lo que para otros fueran pruebas pudieran no serlo para él. La presuposición (2) podría ser verdadera, aunque no lo fuera la presuposición (1), puesto que podría ser que, aceptada una cuestión particular de valores —por ejemplo una cuestión moral—, todo disputante se viera obligado a aceptar como pertinente cualquier prueba, las mismas pruebas que los demás, aunque siempre podría rehusarse a sacar conclusiones morales o a discutir nada en que entraran términos morales. No queremos dar a entender que 'pueda rehusarse a sacar la conclusión', en el sentido trivial según el cual cualquiera puede rehusarse quizá a sacar determinada conclusión; sino que la cuestión está en que cualquier proposición de valor parece siempre ir más allá que cualquier proposición de hecho, de modo que podría tener alguna razón para aceptar las premisas fácticas, pero se rehusara a aceptar la conclusión valuatoria. A aquéllos que razonan de esta manera les parece que esto es así porque se sigue de la implicación práctica que obra en la valoración. Cuando alguien emplea una palabra como 'bueno' en sentido 'valuatorio', aunque no en sentido de 'comillas', parece que compromete su voluntad. De esto ha parecido que se sigue inevitablemente que existe una brecha lógica entre hecho y valor, pues ¿no es distinto decir que una cosa es así, y tener una actitud particular hacia la misma; ver que de una acción se seguirán ciertos efectos, y preocuparse de ello? Desde cualquier punto de vista que se considere la valoración, respecto de su característica esencial —en términos de sentimientos, actitudes, aceptación de imperativos, etc.—, queda siempre el hecho de que con la valoración sobreviene un compromiso en nueva dimensión, lo que no se avala con la mera aceptación de los hechos.

Debo sostener que esta opinión va errada, que se ha colocado en lugar inapropiado la implicación práctica del empleo de los términos morales y que si se trata debidamente desaparece la brecha lógica entre las premisas fácticas y la conclusión moral.

En esta argumentación será proficuo tener como modelo la fuerza práctica o 'guía de la acción' de la palabra 'lesión' (injury) que en muchos aspectos, aunque no en todos, es semejante a la [fuerza] * de los términos morales. Es claro —según veo— que toda lesión necesariamente es algo malo y que, por ende, algo que como tal todo el mundo tiene razón de evitar y que los filósofos sentirán la inclinación a decir que cualquiera que emplee 'lesión' en su sentido pleno de 'guía de la acción' se compromete a evitar las cosas que denomina lesiones. Surgirán entonces las conocidas dificultades respecto del hombre que dice saber que debe hacer algo pero que no lo hará, así como respecto de las debilidades de la voluntad. Supongamos que, en vez [de atender a esas dificultades] *, nos detenemos a considerar qué cosas cuentan como lesiones, a fin de ver si no es aquí donde se inicia su conexión con la voluntad. Como se ha visto, se dice que un hombre ha sido lesionado cuando ha perdido el funcionamiento cabal de una parte de su cuerpo, debido a la lesión. Se sigue que sufre una incapacidad o que está expuesto a sufrirla; con una lesión en la mano no tendrá tanta capacidad para asir las cosas, sujetarlas, unirlas, astillarlas, etc. Si la deficiencia está en los ojos serán otras mil cosas las que no podrá realizar, y en ambos casos diremos que con frecuencia no podrá conseguir lo que quiere o evitar lo que quiere evitar.

Algunos filósofos se asirán a la palabra 'querer' y dirán que si suponemos que alguien quiere las cosas que una lesión en su cuerpo le previene alcanzar, caemos en la presuposición de la 'pre-actitud', y que cualquiera que no quiera tales cosas puede rehusarse a emplear 'lesión' con sentido prescriptivo o de 'guía de la acción'. Y así, puede parecer que la única manera de hacer una conexión *necesaria* entre 'lesión' y las cosas vitandas, es decir, que sólo se emplea en un sentido de 'guía de la acción' cuando se aplica a algo que el hablante intenta evitar. Pero hemos de atender cuidadosamente al movimiento principal de

este argumento y poner entre interrogantes la idea de que alguien no quiera algo para lo que precise de manos y ojos. Manos y ojos, lo mismo que los oídos y las piernas, tienen su papel en tantas operaciones que sólo cabría que un hombre no las quisiera si careciese por completo de necesidades. Que tal gente exista en hospitales psiquiátricos no es del caso; que alguien quiera el uso de sus miembros es algo lógico si quiere conseguir algo.

No entiendo qué pueda tener en mente quien niegue tal proposición. ¿Querrá acaso cambiar los hechos de la existencia humana de forma que con sólo desear, o con el sonido de la voz, se hagan las cosas? ¿O quizá, a lo mejor, está pensando en encuadrar las circunstancias de la existencia de algún individuo que viva en el mundo real, pero imaginando —*v. g.*— que se trata de un príncipe cuyo criado sembrará y recolectará y transportará para él, empleando sus manos y ojos en su servicio, para que no necesite utilizar las propias? Supongamos que fuera posible tal caso; es bárbaramente improbable, pero imaginemos que no lo es. Con todo, es claro que podríamos afirmar que cualquiera tendría razón para evitar lesiones, pues si incluso se pudiera decir que hasta el fin de sus años, por extraña urdimbre de circunstancias, jamás tuvo necesidad de sus ojos o de sus manos, esto no podría haberse previsto; sólo cambiando una vez más los hechos de la existencia humana, y suponiendo que fuera previsible toda vicisitud, se podría hacer tal suposición.

Esto no es negar que una lesión pueda traer más ventaja accidental que daño necesario; basta con pensar cuando había orden de que los aptos para todo servicio debían entrar en filas. En tales casos podría preverse la ventaja de sufrir alguna lesión y preferirla en vez de evitarla. En este sentido la palabra 'lesión' difiere de términos tales como 'injusticia'; la fuerza práctica de 'lesión' significa sólo que alguien tiene *una* razón para evitar las lesiones, no que tenga una razón omnipresente para hacerlo así.

Se advertirá que en esta disertación sobre la fuerza 'guía de la acción' de lesión, se vincula dicha fuerza a razones para actuar y no con el hecho de realizar algo realmente. No creo, sin embargo, que por ello se convierta en menos bueno el patrón de la fuerza 'guía de la acción' de los términos morales. Los filósofos que han supuesto que se requería la acción real, si 'bueno' se decía emplear con valoración sincera, se han topado con dificultades debidas a la debilidad de la voluntad, y sin duda concederán que se hará bastante si se logra demostrar que cualquiera tiene razón para aspirar a la virtud y evitar el vicio. Pero, ¿es sumamente dificultoso esto, si se atiende al tipo de cosas que cuentan como virtud y vicio? Veamos, por ejemplo, las virtudes cardinales: prudencia, justicia, fortaleza y templanza. Es claro que todo el mundo necesita prudencia, pero ¿no necesita también resistir a la tentación del placer cuando le podría sobrevenir un perjuicio? Y ¿cómo se podría sostener que jamás necesitará enfrentarse a algo temible, por causa de algún bien? No es del todo claro qué se querría dar a entender si alguien dijera que la templanza y la fortaleza no eran cualidades buenas y esto no por el sentido 'loable' de los *vocablos*, sino por lo que son la fortaleza y la templanza.

Quiero emplear estos ejemplos para mostrar la artificialidad de las nociones de 'comendación' y de 'proactitudes', según se las emplea comúnmente. Quienes hablan de tales cosas dirán que, luego de haber aceptado los hechos —como que X es la clase de hombre que escala un monte peligroso o que se enfrenta a un patrón irascible para pedirle mayor paga y, en general, que encara lo temible por algo que lo merezca—, quedará la cuestión de la 'comendación' y de la 'valoración'. Si se trata de la palabra 'fortaleza', preguntarán si quien hable de otro como fuerte se supone que lo comienda o no. Si decimos que sí, insistirán en que el juicio sobre la fortaleza *va más allá de los hechos* y por lo mismo puede ser desechado por alguien que no quiera actuar así; si decimos que

no, alegarán que 'fortaleza' se emplea en un sentido puramente descriptivo o 'entrecomillado' y que no tenemos un ejemplo del empleo valoratorio del lenguaje, que es la incumbencia especial de los filósofos morales. ¿Qué sentido, pues, tiene la cuestión, 'está comendando'? ¿Cuál es ese elemento extra que se supone está presente o ausente una vez se han determinado los hechos? No se trata de que complazca el hombre de fortaleza, o de considerarlo del todo bueno, sino de 'comendarlo por su fortaleza'. ¿Cómo se podrá hacer? La respuesta que se dará es que sólo comendamos a alguien, si hablamos de él como valiente, cuando nosotros aceptamos el imperativo 'tengo que ser valiente' al tratarse de nosotros. Pero esto es del todo innecesario. Puedo hablar de alguien como que tiene la virtud de la fortaleza y reconocer a ésta como virtud en el sentido propio, sin ignorar que soy completamente cobarde y que no tomo resoluciones para reformarme. Puedo saber que sería mejor si fuera valiente, pero puedo también saber que jamás haré nada en pro de tal cosa.

Si alguien dijera que la fortaleza no es una virtud, estaría diciendo que no es una cualidad con la que alguien actuara bien. Quizá estaría pensando en que un individuo puede ser peor debido precisamente a su valor, lo que es cierto, pero sólo porque puede ocurrir un perjuicio accidental. Por ejemplo, una persona denodada puede haber subestimado un peligro y correr hacia un desastre, que el hombre amilanado habría evitado porque no habría estado dispuesto a afrontar riesgo alguno. Así, su valor —al igual que cualquier otra virtud— podría ser causa de su daño, porque al tenerlo cayera en algún orgullo desastroso [4]. De manera semejante, quienes ponen en tela de juicio la virtud de la templanza piensan probablemente no en la virtud en sí, sino en hombres cuya templanza ha consistido en resistir al placer en aras

[4] Comparar Tomás de Aquino, *Summa Theologica*, I-II, q. 55, Art. 4.

de algún bien ilusorio o en aquellos que han hecho su orgullo de esta virtud.

Pero, se preguntará, ¿y la justicia? Pues mientras la prudencia, el valor y la templanza son cualidades que benefician a quien las posee, la justicia diríase que beneficia a los demás y que resulta en menoscabo del hombre justo. La justicia, cual aquí se trata, o sea como virtud cardinal, se refiere a todas aquellas cosas que se deben a los demás: es bajo la injusticia cuando reinan el crimen, el robo y la mentira, o el retener lo que los padres deben a los hijos, o viceversa, o los tratos que en lenguaje común se llaman injustos. Así, el hombre que evita la injusticia se hallará falto de las cosas que ha devuelto a su dueño, sintiéndose incapaz, además, de sacar ventajas engañando o mintiendo, y con todas aquellas dificultades que Trasímaco pinta en el primer libro de la República para demostrar que la injusticia es más ventajosa que la justicia, si la persona tiene fuerza e ingenio. Se nos preguntará ahora cómo, según nuetra teoría, la justicia puede ser una virtud, y un vicio la injusticia, pues va a ser difícil demostrar que alguien necesite ser justo, como necesita de sus ojos y de sus manos, o necesita la prudencia, la fortaleza y la templanza.

Antes de responder a este interogante tengo que establecer que si no se puede contestar, entonces la justicia no puede ser recomendada como virtud. El quid de esto no es demostrar que ha de tener respuesta, dado que la justicia es virtud, sino más bien señalar que debemos siquiera considerar la posibilidad de que la justicia no fuera virtud. Este prenctando fue tomado seriamente por Sócrates en la *República*, puesto que todos presumieron que si Trasímaco llegaba a demostrar su premisa —sobre que la injusticia era más ventajosa que la justicia— se seguiría la conclusión de que alguien con fuerzas para salirse con la suya mediante la injusticia, tendría razón para seguir por ese camino como el mejor para él. Es hecho sorprendente de la filosofía moral mo-

derna que nadie vea dificultad en admitir la premisa de Trasímaco rechazando su conclusión, debido a que la posición de Nietzsche en este punto está más próxima de la de Platón como lejano está éste de los moralistas académicos de nuestro tiempo.

En la *República* se supone que si la justicia no es un bien para el justo, los moralistas que la recomiendan como virtud están perpetrando un fraude. Si asiento a esto se me preguntará dónde exactamente entra el fraude, dónde se dice la mentira de que la justicia sea provechosa para el individuo. Como respuesta preliminar podemos preguntar, ¿cuánta gente está dispuesta a confesar francamente que la injusticia es más proficua que la justicia? Dejando de lado, como lo hemos hecho en todo este artículo, las creencias religiosas, que complicarían la cuestión, supondremos que una persona duramente atea hubiera preguntado, '¿por qué he de ser justo?' (Quienes crean que hay algo que no está bien en la pregunta, pueden emplear su artimaña favorita de cribar el 'significado valorativo' imaginando que la pregunta es '¿Por qué he de ser «justo»?'). Si le replicáramos: 'En cuanto a usted se refiere le irá mejor si es injusto, pero a nosotros nos conviene más que sea justo, por lo que intentaremos que lo sea', es probable que se dedicara a enterarse de qué pie cojeábamos y procurara no ser atrapado; por lo demás, no creo que quienes opinan que no es preciso demostrar que la justicia es provechosa para el hombre justo acepten que no hay más que decir.

La cuestión palpitante es: '¿Podemos dar a alguien, fuerte o débil, una razón por la que tenga que ser justo?' No vale escabullirse diciendo que, pues 'justo' e 'injusto' son 'palabras guías de la acción', no se puede siquiera preguntar '¿Por qué he de ser justo?'. Enfrentado al argumento, quien desee ser injusto no tiene más que cuidar de evitar la *palabra*, pues no se le ha dado razón alguna de por qué no ha de hacer las cosas que otros llaman 'injustas'. Se dirá probablemente que se le ha dado una razón, hasta donde

se puede dar una razón para hacer o dejar de hacer algo, pues la cadena de razones ha de concluir en algún punto, ya que parece que alguien pueda siempre rechazar una razón que otro aceptaría. Mas esto es una equivocación. Hay respuestas a la pregunta '¿por qué he de?' que finiquitan la cuestión, mientras que otras no lo hacen. Hume demostró cómo una respuesta cerraba la serie, en el siguiente pasaje:

'Si le preguntas a alguien *por qué hace ejercicio*, responderá *porque desea conservar la salud*. Si luego le preguntas, *por qué desea tener salud*, replicaría inmediatamente: *porque la enfermedad es dolorosa*. Si todavía sigue preguntándole *por qué no quiere el dolor*, es imposible que pueda dar respuesta. Se trata de un final que no se puede referir a otro objeto.' (*Enquiries*, Apéndice I, V.) Hume podría haber concluido la serie con hastío: la enfermedad trae consigo hastío y nadie tiene por qué dar una razón de por qué no quiere ser molestado, de la misma manera que no tiene que dar respuesta de por qué busca lo que le interesa. En general, todo el mundo recibe una razón para actuar cuando se le muestra una senda para llegar a algo que desea, pero hay deseos para los que tiene sentido la pregunta '¿por qué deseas esto?', no así para otros [5]. Parece claro que en esta división la justicia cae en el lado opuesto del placer, del interés y de cosas semejantes. '¿Por qué no he de hacer esto?' no se responde con las palabras 'porque es injusto', como se puede responder mostrando que la acción acarreará hastío, soledad, dolor, displicencia o alguna incapacidad; por esto no es verdad decir que 'es injusto' da razón en tanto puede darla. 'Es injusto' da razón sólo si se puede demostrar que la naturaleza de la injusticia es tal que se enlaza necesariamente con lo que alguien desea.

Esto muestra por qué la cuestión de si la justicia es buena o no para el hombre justo trae cola y por

5 Para una discusión excelente sobre los motivos para actuar, véase G. E. Anscombe, *Intention*, § 34-40.

qué quienes aceptan la premisa de Trasímaco y rechazan su conclusión se sitúan en posición dudosa. Recomiendan la justicia para cada uno, como algo que el individuo tiene razón para seguir, pero cuando se les reclama que muestren por qué se ha de obrar así no siempre pueden responder. Esta última aserción no depende de ninguna 'teoría egoísta de la naturaleza humana' en sentido filosófico. Es posible con frecuencia darle a alguien una razón de por qué ha de obrar de alguna manera, mostrándole que otra persona puede sufrir si él no actúa de ese modo; el bien de otro puede serle realmente de más prez que el propio. Pero el afecto que las madres sienten por sus hijos, los amantes uno por otro y los amigos entre sí, no nos llevará lejos cuando se nos pregunte por qué una persona tiene que ser justa; en parte porque no se extiende muy lejos y en parte porque las acciones dictadas por benevolencia y por la justicia no siempre son las mismas. Supongamos que debo dinero a alguien. '... ¿y si es mi enemigo y me da motivos para que lo odie?, ¿y si es un hombre malvado que merece el desprecio de toda la humanidad?, ¿y si es un avaro que ningún provecho sacará de lo que le devuelva?, ¿y si es un pródigo empedernido que recibirá más daño que provecho de tener mucho?' [6]. Incluso si la práctica general de la justicia pudiera reducirse al motivo de la benevolencia universal —deseo de la mayor felicidad posible para el mayor número también posible— mucha gente habría que no tendría interés en ello. Así, pues, si la justicia es algo que se recomienda por los anteriores motivos, miles de caracteres difíciles dirán que no se les ha dado razón alguna para practicarla, y lo mismo dirían muchos más si no fueran o demasiado tímidos o demasiado estúpidos para interrogar acerca del código de conducta que se les ha enseñado desde siempre. Así, pues, dada la premisa de Trasímaco, su punto de vista es razonable; no tenemos razón alguna par-

6 Hume, *Treatise*, Libro III, Parte II, Secc. 1.

ticular para admirar a aquéllos que practican la justicia por timidez o por estulticia.

Me parece, por tanto, que si se acepta la tesis de Trasímaco, las cosas ya no pueden ser como antes. Tendremos que admitir que la creencia sobre la que se fundaba el *status* de la justicia como virtud está equivocada, y si deseamos que la gente sea justa tenemos que recomendarle la justicia de otra manera. Tendremos que admitir que la injusticia es más provechosa que la justicia, al menos para los fuertes, y entonces hacer lo que podamos para demostrar que es difícil que alguien salga limpio de polvo y paja siendo injusto. Nos queda, es claro, la alternativa de no movernos, en espera de que la gente en su mayoría seguirá lo convencional respecto a algún tipo de justicia y no hará preguntas raras; pero este procedimiento puede quedar contagiado de cierto escepticismo, incluso entre aquéllos que no saben a ciencia cierta qué es lo que no anda bien; quedaríamos también a merced de cualquiera que fuera capaz y quisiera sacar al sol nuestro fraude.

¿Es cierto, sin embargo, que no es la justicia lo que el hombre requiere en sus tratos con sus prójimos, puesto caso que sea fuerte? A quienes creen que pueden salirse con la suya perfectamente siendo injustos, se les debería rogar que dijeran con exactitud cuánto tiempo puede llegar a vivir un hombre. Sabemos que ha de practicar la injusticia siempre que el acto injusto le reporte ventaja. Pero, ¿qué dirá él?, ha de confesar que no reconoce derechos a los demás, o está fingiendo? En el primer caso, incluso aquéllos que se confabulan con él, han de saber que si cambia la fortuna o se altera su afecto procurará estafarlos, y que él está tan alerta sobre su traición como ellos lo están respecto de la de él. Quizá se imaginan al injusto feliz, como sucede en el Libro II de la *República,* cual un mentiroso muy astuto y como un actor que sabe combinar la injusticia completa con la apariencia de justicia: está dispuesto a tratar a los demás sin piedad alguna, pero finge que no hay

cosa que menos diga con él. Los filósofos hablan con frecuencia como si algún individuo pudiera encubrirse frente a los que le rodean, pero tal presuposición es dudosa, y en todo caso el precio que debería pagar en vigilancia sería colosal. Si dejara que siquiera algunos supieran de sus mañas, debería guardarse de ellos; si a nadie comunica el secreto, ha de estar en continua circunspección, para evitar que alguna espontaneidad lo delate. Estos hechos son importantes porque la necesidad que el hombre tiene de obrar justamente con los demás depende del hecho de que éstos son hombres y no objetos inanimados o bestias. Si alguien sólo necesita a los demás como puede necesitar los enseres domésticos, y si los hombres pudieran manipularse como enseres, o se les pudiera golpear para que se sometieran como si fueran asnos, otro sería el caso. Pero cual están las cosas, la suposición de que la injusticia es más provechosa que la justicia es dudosa, aunque, como la cobardía y la intemperancia, accidentalmente puede resultar ventajosa.

La razón de por qué a cierta gente parece tan imposiblemente difícil demostrar que la justicia es más provechosa que la injusticia está en que se consideran aisladamente actos justos particulares. Es del todo cierto que si un hombre es justo se sigue que estará dispuesto, aun en el caso de circunstancias muy adversas, a afrontar incluso la muerte antes de ser injusto, por ejemplo, permitiendo que un inocente pague por un crimen que no ha cometido. Para él, su justicia le reporta desventajas y, no obstante, como cualquier otro, tiene buena razón para ser justo y no injusto. Podría haber echado mano de las dos cosas y mientras poseía la virtud de la justicia, estar dispuesto a ser injusto si se terciara alguna gran ventaja. Quien tiene la virtud de la justicia no está dispuesto a hacer ciertas cosas, y si resulta que se presta fácilmente a la tentación, veremos que después de todo sí estaba dispuesto.

VII

COMO DERIVAR 'DEBE' DE 'ES' [1]

John R. Searle

De *Philosophical Review*, Vol. 73 (1964), pp. 43-58. Reimpreso con la venia del autor y de *Philosophical Reivew*.

I

Se dice a menudo que no es posible derivar 'debe' de 'es'. Esta tesis, que procede de un famoso pasaje del *Treatise* de Hume, aunque no es tan clara como podría serlo, lo es al menos en un sentido lato: hay una clase de proposiciones de hecho que es distinta lógicamente de la clase de las proposiciones de valor. Ningún conjunto de proposiciones de hecho contiene por sí mismo proposiciones de valor. Y dicho con terminología más contemporánea, ningún conjunto

1 Ante el Stanford Philosophy Colloquim y la Pacific Division of the American Fhilosophical Association se leyeron versiones anteriores de este mismo artículo. Debo agradecer a mucha gente sus comentarios y críticas proficuos, especialmente a Hans Herzberger, Arnold Kaufmann, Benson Mates, A. L. Melden y Dagmar Searle.
[Este artículo ha sido muy discutido. Véase, *v. g.*, J. y J. Thomson, 'How not to Derive «Ought» from «Is»», *Philosophical Review*, (1964); también A. Flew y otros en *Analysis*, de 1964 a 1966. Tocan este tema J. Searle, 'Meaning and Speech Acts', *Philosophical Review* (1962) y a la contribución de J. Searle ('What is a Speech Act?') en *Philosophy in America*, a cargo de M. Black (George Allen and Unwin, Londres, 1965). Ed.]

de proposiciones *descriptivas* puede contener proposición alguna *valorativa*, a menos que se le añada siquiera una premisa valorativa. Creer lo contrario es cometer lo que se ha llamado la falacia naturalista.

Trataré de demostrar un ejemplo opuesto a esta tesis[2]. No se ha de suponer, es claro, que un solo ejemplo contrario puede refutar una tesis filosófica, pero en el caso presente, si podemos presentar un ejemplo contrario y, además, dar alguna razón o explicación de cómo y por qué es ejemplo en contra, ofreciendo para mayor abundamiento una teoría que respalde nuestro ejemplo contrario —teoría que generará infinito número de ejemplos en contra—, podremos como mínimo echar considerable luz sobre la tesis original y, posiblemente, si logramos todas esas cosas, podremos inclinarnos incluso hacia el punto de vista de que el propósito de esa tesis era más restringido que cuanto habíamos supuesto originalmente. El ejemplo en contrario procederá tomando una proposición o proposiciones que cualquier defensor de la tesis aceptaría, si fuesen puramente fácticas o 'descriptivas' (no es necesario que aparezca realmente la palabra 'es'), y así demostrar cómo están relacionadas lógicamente con otra proposición que el defensor de la tesis consideraría como claramente 'valoratoria'. (En el caso que presento contendrá un 'debe'.)[3].

Considérese la siguiente serie de proposiciones:

(1) Ticio profirió las palabras 'Con esto te prometo, Cayo, pagarte cinco dólares'.

(2) Ticio prometió pagar a Cayo cinco dólares.

(3) Ticio se puso bajo (asumió) la obligación de pagar a Cayo cinco dólares.

[2] En su versión moderna. No me refiero a la manera como Hume maneja este problema.

[3] Si se logra esto, habremos cubierto la brecha entre lo 'valorativo' y lo 'descriptivo' y, por ende, habremos demostrado que existe debilidad en esa misma terminología. Por ahora, sin embargo, mi estrategia es mantener la terminología, presumiendo que las nociones de valorativo y descriptivo son bastante claras. Al final del escrito declararé en qué respectos contienen confusión.

(4) Ticio está bajo la obligación de pagar cinco dólares a Cayo.

(5) Ticio debe pagar cinco dólares a Cayo.

Defenderé respecto de esta lista que la relación entre una proposición y la siguiente, si bien no es en todos los casos de 'implicación', sin embargo no es sólo una relación contingente, y que la proposición adicional, necesaria para convertir la relación en una de implicación, no es preciso que contenga proposiciones valorativas, principios morales o cosas por el estilo.

Empecemos. ¿Cómo se relaciona (1) con (2)? En algunas circunstancias, la expresión de las palabras entre comillas de (1) equivale al acto de hacer una promesa. Y es parte o consecuencia del significado de las palabras en (1) que en tales circunstancias el pronunciarlas es prometer. 'Con esto te prometo' es un paradigma para ejecutar el acto de prometer descrito en (2).

Estipulemos este hecho paradigmático en forma de una premisa extra:

(1a) Bajo ciertas condiciones *C*, quienquiera que profiera las palabras (proposición) 'Con esto te prometo pagarte, Cayo, cinco dólares', promete pagar a Cayo cinco dólares.

¿Qué tipo de cosas se contiene bajo la rúbrica 'condiciones *C*'? Lo que va contenido es todas esas condiciones, esos asuntos, que son necesarias y suficientes condiciones para la pronunciación de las palabras (proposición), de forma que constituyan la ejecución válida del acto de prometer. Las condiciones incluirán cosas como que el hablante está en presencia del oyente Cayo, que ambos están en estado consciente, que ambos hablan el mismo idioma, que están hablando en serio, que el hablante sabe lo que está haciendo, que no está bajo la influencia de drogas, ni hipnotizado, ni representando en el teatro, ni contando un chiste o relatando un suceso, etc. La lista tendrá que ser algo inacabable, puesto que los límites del concepto de promesa, como los límites de la mayoría

de los conceptos del lenguaje natural, son algo laxos[4]. Pero una cosa está clara; por laxos que puedan ser esos límites y por difícil que resulte decidirse en los casos marginales, las condiciones bajo las cuales puede afirmarse correctamente que está haciendo una promesa quien profiere 'Con esto te prometo', son condiciones indudablemente empíricas.

Añadamos como premisa extra la suposición de que se cumplen esas condiciones.

(1b) Las condiciones *C* han lugar.

De (1), (1a) y (1b), derivamos (2). El argumento tiene esta forma: Si *C*, entonces (si *D* entonces *P*), donde *C* son las condiciones, *D* la declaración que se profiere y *P* la promesa. Añadiendo las premisas *D* y *C* a esta hipótesis, derivamos (2). Y, hasta donde me es dado ver, no amaga ninguna premisa moral en este rimero lógico. Es preciso decir más acerca de la relación de (1) con (2), pero lo reservo para después.

¿Qué relación existe entre (2) y (3)? Presumo que, por definición, prometer es situarse bajo una obligación. No será completo ningún análisis del concepto de prometer si no contiene la característica de que el promisor se somete, acepta, reconoce o se coloca bajo obligación frente al depositario de efectuar alguna acción futura, de ordinario para beneficio del depositario de la promesa. Alguien puede sentirse tentado a pensar que el prometer se puede analizar en términos de que origina expectaciones en los oyentes de uno, o en los que hagan sus veces, pero con un poco de reflexión se verá que la distinción fundamental entre proposiciones de intención, por un lado, y promesas, por el otro, estriba en la naturaleza y grado de compromiso u obligación asumida al prometer.

Por lo tanto, me siento inclinado a decir que (2) implica ineludiblemente (3), pero no tengo inconve-

[4] Además, el concepto de promesa es miembro de una clase de conceptos que adolecen de cierta laxitud de especial tipo, a saber, de anulabilidad. Cf. H. L. A. Hart 'The Ascription of Responsibility and Rights', *Logic and Language*. Primera serie a cargo de A. Flew (Oxford, 1951).

niente en que se añada —por razones de aliño formal— la premisa tautológica:

(2a) Todas las promesas son actos de situarse bajo (asumir) la obligación de hacer la cosa prometida.

¿Qué relación hay entre (3) y (4)? Si alguien se ha situado bajo una obligación, entonces, siendo iguales las demás cosas, se está bajo una obligación. Esto, a mi modo de ver, es también una tautología. Desde luego, es posible que ocurran cosas que desliguen a uno de la obligación asumida y de ahí la necesidad de la caución *caeteris paribus*. Para que haya implicación entre (3) y (4), por ende, necesitamos una proposición cualificante que la efectúe:

(3a) Las demás cosas son iguales.

Los formalistas, al igual como ocurrió en el paso de (3) a (4) desearán añadir la premisa tautológica:

(3b) Todos aquéllos que se sitúan bajo una obligación están, siendo iguales las demás cosas, bajo una obligación.

El paso de (3) a (4) es, pues, de la misma forma que el paso de (1) a (2): Si *I*, entonces (si *SBO*, entonces *BO*), donde *I* equivale a siendo iguales las demás cosas; *SBO*, a situarse bajo obligación, y *BO*, a bajo obligación. Conjuntando las dos premisas *I* y *SBO*, derivamos *BO*.

¿Es (3a), la cláusula del *caeteris paribus*, una premisa valorativa larvada? Sin duda parece como si lo fuera, especialmente según la formulación que le he dado, pero creo poder demostrar que, si con frecuencia hay consideraciones valorativas en las preguntas sobre si son iguales las demás cosas, no es lógicamente necesario que tenga que ser así en todos los casos. Pospondré discutir esto hasta el próximo paso.

¿Qué relación existe entre (4) y (5)? Hay aquí una tautología, análoga a la que explica la relación entre (3) y (4), sobre que tiene que hacerse aquello bajo cuya obligación de hacer se está. Y aquí, como en el caso anterior, necesitamos una premisa de la forma:

(4a) Las demás cosas son iguales.

Precisamos de la cláusula *caeteris paribus* para eliminar la posibilidad de que pueda interferir algo extraño a la relación de 'obligación' y 'debe'[5]. Aquí, como en los dos pasos previos, eliminamos la apariencia de entimema, señalando que la premisa al parecer elidida es tautológica, y que, por tanto, si bien formalmente es adecuada, está de más. Si, no obstante, queremos plantear el argumento de manera formal, poseerá la misma forma que el paso de (3) a (4): Si *I*, entonces (si *BO*, entonces *D*), donde *I* equivale a las demás cosas quedando iguales; *BO*, a bajo obligación, y *D*, a debe. Conjuntando las premisas *I* y *BO*, derivamos *D*.

Ahora diré algo de la frase 'siendo iguales las demás cosas' y de cómo funciona en la derivación que he intentado. Este tema, y el de la anulabilidad, estrechamente emparentado con él, son dificultosos en extremo y no intentaré hacer otra cosa que justificar mi alegato de que la satisfacción de la condición no implica necesariamente algo valoratorio. La fuerza de la expresión 'siendo iguales las demás cosas', en la contingencia presente, es más o menos ésta. A menos que tengamos alguna razón (o sea, salvo que estemos dispuestos realmente a dar alguna razón) para suponer que la obligación está anulada (Paso 4), o que el agente no tiene que cumplir la promesa (Paso 5), entonces la obligación se mantiene en pie y se debe cumplir la promesa. No es parte de la fuerza de la frase 'siendo iguales las demás cosas' que, con el fin de satisfacerla, tengamos que asentar una proposición negativa universal que declare que jamás podría darse razón alguna para suponer que el agente no está bajo la obligación, o no deba, cumplir lo prometido.

5 El *caeteribus paribus* de este paso excluye tipos de casos algo distintos de los excluidos en el paso anterior. En general decimos 'Asumió una obligación, pero sin embargo no está (ahora) bajo obligación', cuando se ha *removido* dicha obligación'. Mas decimos 'Está bajo obligación, pero no la ha de satisfacer', en casos en que la obligación queda *cancelada* por otras consideraciones, v. g., otra obligación que tiene prioridad.

Basta para satisfacer la condición que no se pueda dar de hecho razón alguna en contrario.

Si se da alguna razón para suponer que la obligación está anulada o que aquel que promete no debe cumplir una promesa, entonces surge, de manera característica, una situación que apela a la valoración. Supongamos, por caso, que el acto prometido es indebido, pero estamos de acuerdo en que el de la promesa asumió una obligación. ¿Debe guardar su promesa? No hay procedimiento estatuido para decidir tales casos por adelantado, y se impone una valoración (si ésta es en realidad la palabra). Pero, salvo que poseamos alguna razón en contrario, se satisface la condición del *caeteris paribus* y no se necesita valoración alguna, quedando resuelta la cuestión de si debe hacer algo diciendo 'prometió'. Queda siempre abierta la posibilidad de que hayamos de hacer una valoración para derivar 'debe' de 'prometió', porque tengamos que sopesar un argumento en contrario. Pero la valoración no es lógicamente necesaria en todos los casos, pues podría ser que, de hecho, no se presentaran argumentos en contra. Me siento inclinado a pensar, por tanto, que no hay nada que sea necesariamente valoratorio respecto del *caeteris paribus*, aun cuando el decidir si se ha satisfecho esa condición a menudo exija valoraciones.

Pero supongamos que ando equivocado en esto. ¿Nos salvaríamos de la creencia de que existe un abismo lógico inabarcable entre el 'es' y el 'debe'? No lo creo, pues siempre nos quedaría poder enmendar mis pasos (4) y (5) de manera que incluyeran la cláusula del *caeteris paribus* como parte de la conclusión. Así, de nuestras premisas habríamos derivado 'Siendo iguales las demás cosas, Ticio tiene que pagar cinco dólares a Cayo', y esto sería suficiente para refutar la tradición, pues habríamos mostrado que existe una relación de implicación entre las proposiciones descriptivas y las valorativas. No fue el hecho de que las circunstancias extremas pueden hacer que las obligaciones se vuelvan írritas lo que

condujo a los filósofos a la falacia de la falacia naturista, sino más bien una teoría del lenguaje, como veremos posteriormente.

Así, pues, hemos derivado (en un sentido tan estricto de 'derivar' como cabe en los lenguajes naturales) un 'debe' de un 'es'. Y las premisas extra que se han precisado para que funcionara la derivación no fueron por ninguna causa ni morales ni valoratorias por naturaleza; consistieron en presunciones y tautologías empíricas y en descripciones del empleo de las palabras. Se ha de señalar también que el 'debe' es un debe 'categórico', no 'hipotético'. (5) no dice que Ticio debe pagar si desea tal y tal cosa; dice que ha de pagar, y punto. Nótese también que los pasos de la derivación se llevan a cabo en tercera persona. No concluimos 'debo' de 'dije que «yo prometía»', sino 'debe' de 'dijo que «yo prometo»'.

La prueba explana la conexión existente entre la declaración de ciertas palabras y el acto locutorio de prometer y, luego, a su vez, lleva la promisión a la obligación y se mueve de la obligación al 'debe'. El paso de (1) a (2) es distinto radicalmente de los otros y requiere comentario especial. En (1) construimos 'prometo con esto...' como una frase consagrada, con determinado significado. Es consecuencia de ese significado que la pronunciación de esa frase bajo ciertas condiciones sea el acto de prometer. Así, al presentar las expresiones citadas en (1) y al describir su empleo en (1a) es como si hubiéramos invocado la institución de la promisión. Podríamos haber empezado con una premisa todavía más a ras del suelo que (1), diciendo:

(1b) Ticio profirió la secuencia fonética: 'Con ésto / te prométo // Cáyo // pagárte / sinco dólares' ///

Entonces habríamos requerido premisas extra, empíricas, declarando que esa secuencia fonética iba unida de determinadas maneras con determinadas unidades significativas pertenecientes a determinados dialectos.

Los pasos de (2) a (5) son fáciles relativamente. Nos apoyamos en conexiones definitorias entre 'prometer', 'obligar' y 'debe', pero el único problema que surge es que las obligaciones pueden quedar anuladas o supeditadas por distintas causas, y esto lo hemos de tomar en cuenta. Resolvemos nuestra dificultad añadiendo aún más premisas, para dejar en claro que no existen prenotandos en contra y que las demás cosas quedan iguales.

II

En esta sección deseo discutir tres objeciones posibles a la derivación.

Primera objeción

Como la primera premisa es descriptiva y la conclusión valorativa, tiene que haber una premisa valorativa larvada en la descripción de las condiciones de (1b).

Hasta aquí, esta acotación no hace más que pedir la cuestión, pues supone la existencia de una brecha lógica entre lo descriptivo y lo valorativo que la derivación ha de cuestionar. Para que la objeción valga, su defensor debería mostrar exactamente cómo (1b) ha de contener una premisa valorativa y qué tipo de premisa puede ser. La pronunciación de ciertas palabras en ciertas condiciones sin más *es* prometer y la descripción de esas condiciones no precisa de ningún elemento valorativo. Lo esencial es que en la transición de (1) a (2) nos movemos de una especificación de cierto enunciado de palabras a la especificación de cierto acto locutorio. Se logra el paso porque el acto locutorio es convencional, y la pronun-

ciación de las palabras de acuerdo con los convencionalismos es lo que constituye precisamente la ejecución de ese acto locutorio.

Variante de esta primera objeción es decir: todo lo que se ha demostrado es que 'promesa' es concepto valoratorio, no descriptivo. Pero esta apostilla es también una *petitio quaestionis* y al cabo resultará desastrosa respecto de la distinción original entre descriptivo y valorativo. Pues que alguien pronuncie ciertas palabras y que estas palabras tengan el significado que tienen son sin duda actos objetivos. Y si la declaración de estos dos actos objetivos más la descripción de las condiciones del enunciado son suficientes para implicar la proposición (2), que el objetor sostiene que es una proposición valorativa (Ticio prometió pagar cinco dólares a Cayo), entonces se deriva una conclusión valorativa de premisas descriptivas, sin siquiera pasar por (3), (4) y (5).

Segunda objeción

Finalmente, la derivación estriba en el principio de que se deben cumplir las promesas, y éste es un principio moral; por ende, valorativo.

Yo no sé si 'se deben cumplir las promesas' es un principio 'moral', pero séalo o no, también es tautológico, pues no es más que una derivación de dos tautologías:

Todas las promesas son (crean, son asunciones de, son aceptaciones de) obligaciones

y

se deben cumplir (guardar) las obligaciones.

Lo que se ha de explicar es por qué ha habido tantos filósofos que no han logrado ver el carácter tautológico de este principio. Creo que han sido tres las cosas que han impedido percatarse de ello.

La primera es la falla para distinguir cuestiones externas relativas a la institución de la promisión, de las cuestiones internas que se plantean dentro del

marco de la institución. Las preguntas '¿Por qué poseemos una institución como el prometer?' y '¿Debemos tener tales formas institucionalizadas de obligación como la promisión?' son externas y se plantean en torno, no dentro de la institución del prometer. Y la pregunta '¿Se han de cumplir las promesas?', se puede confundir con, o se puede tomar como (y creo que con frecuencia se ha tomado como) una pregunta externa expresable más o menos así: '¿Se ha de aceptar la institución del prometer?' Pero, tomada literalmente como pregunta interna, como una pregunta acerca de las promesas y no acerca de la institución del prometer, la pregunta '¿Se han de cumplir las promesas?' es tan hueca como la interrogación '¿Tienen tres lados los triángulos?' Reconocer algo como promesa es conceder que, siendo iguales las demás cosas, se ha de cumplir.

Un segundo hecho que ha obnubilado la cuestión es éste. Hay muchas situaciones, tanto reales como imaginarias, en que no se ha de cumplir la promesa, en que la obligación de cumplir una promesa queda contrarrestada por consideraciones ulteriores, y fue por esta razón por lo que necesitamos del engorroso *caeteris paribus* en nuestra derivación. Pero el hecho de que las obligaciones puedan quedar supeditadas no muestra que no las hubiera anteriormente. Al contrario; y son esas obligaciones originales las que bastan para hacer válida la prueba.

Hay, con todo, un tercer factor, que es el siguiente. Muchos filósofos no logran ver todavía la fuerza total que hace de decir 'Con esto prometo' una expresión ejecutoria. Al proferirla, uno ejecuta, mas no describe, el acto de prometer. Si la promisión se considera como un acto locutorio de clase diferente al describir, entonces es más fácil ver que una de las características del acto es la asunción de una obligación. Pero si se piensa que la expresión 'yo prometo' o 'con esto prometo' es un tipo particular de descripción —por ejemplo, del estado mental de uno—

entonces la relación entre prometer y obligación ha de parecer muy misteriosa.

Tercera objeción

En la derivación hemos echado mano sólo de un sentido fáctico o 'entrecomillado' de los términos valorativos empleados. Por ejemplo, un antropólogo que observara el comportamiento de los anglosajones podría seguir esas derivaciones, sin incluir nada valorativo. Así, el paso (2) equivale a 'Hizo lo que se llama prometer' y el paso (5) equivale a 'Según lo cual debería pagar cinco dólares a Cayo'. Pero puesto que todos los pasos de (2) a (5) están en *oratio obliqua* y, por tanto, son proposiciones de hecho disfrazadas, queda sin afectarse la distinción hecho-valor.

Esta objeción no perjudica la derivación, pues lo que dice es sólo que se *pueden* reconstruir los pasos en *oratio obliqua;* que los podemos construir como una serie de proposiciones externas; que podemos construir una prueba paralela (o relacionándola de alguna manera) respecto del habla a que se hace referencia. Pero lo que propugno es que, tomada literalmente, sin adiciones o interpretaciones de *oratio obliqua,* la derivación es válida. Que sea posible construir un argumento similar que no lograra refutar la distinción hecho-valor no demuestra que esta prueba deje de refutarla; en realidad es algo que no roza este asunto.

III

Hasta aquí he presentado un ejemplo en contrario respecto de la tesis de que no se puede derivar 'debe' de 'es', y he considerado tres objeciones posibles. Aun

suponiendo que lo dicho hasta aquí sea verdad, con todo, se siente cierta intranquilidad. Se siente como si en algún lugar hubiera trampa. Podríamos declarar así nuestra intranquilidad: ¿Cómo concediendo un simple hecho respecto de alguien, como el hecho de que haya proferido ciertas palabras o que haya prometido, *me* puedo obligar a admitir que *él* deba hacer algo? Quiero dilucidar brevemente ahora esta más lata consecuencia filosófica que pudiera tener mi derivación, con el intento de obtener los rasgos de la respuesta a esta cuestión.

Empezaré discutiendo los motivos para suponer que no se puede responder en modo alguno.

La inclinación a aceptar una distinción rígida entre 'es' y 'debe', entre descriptivo y valorativo, descansa sobre cierta idea de cómo las palabras se relacionan con el mundo. Es una idea muy atractiva, tanto (para mí al menos) que no es del todo claro hasta qué punto la mera presentación de ejemplos en contra pueda hacerla tambalear. Lo que se requiere es la explicación de cómo y por qué este cuadro empirista clásico no da en el clavo con esos ejemplos en contrario. Para decirlo más brevemente, el cuadro tiene más o menos esta estructura: en primer lugar, presentamos ejemplos de las proposiciones llamadas descriptivas ('mi coche va a 60 km/h', 'Juan mide 1,70', 'Pedro tiene pelo castaño') y los contraponemos a proposiciones llamadas valorativas ('mi coche es bueno', 'Juan tiene que pagar a Pedro cinco dólares', 'Pedro es un hombre fastidioso'). Cualquiera ve que son diferentes. Estipulamos la diferencia diciendo que cuando se trata de proposiciones descriptivas, la cuestión de la verdad o falsedad es decidible objetivamente, pues si sabemos el significado de las expresiones descriptivas, sabemos bajo qué condiciones ratificables objetivamente serán ciertas o falsas. Pero en el caso de las proposiciones valorativas, la situación es muy otra. Saber el significado de las expresiones valorativas no basta de por sí para saber bajo qué condiciones las proposiciones que las contienen

son verdaderas o falsas, porque el significado de las expresiones es tal que las proposiciones no son capaces en absoluto de verdad o falsedad objetivas o fácticas. Cualquier justificación que el hablante pueda dar de alguna de sus proposiciones valorativas importa esencialmente alguna referencia a actitudes que mantiene, a criterios de calificación por él adoptados, o a principios morales a tenor de los cuales ha elegido vivir y juzgar a los otros. Así, pues, las proposiciones descriptivas son objetivas, mientras que las valorativas son subjetivas, y la diferencia es consecuencia de los distintos tipos de vocablos empleados.

La razón subyacente respecto de estas diferencias es que las proposiciones valorativas efectúan cometido por completo diferente al de las proposiciones descriptivas. Su función no es describir características del mundo, sino expresar las emociones del hablante, sus actitudes, alabar o condenar, encomiar o vilipendiar, comendar, recomendar, conminar, etc. Si advertimos los diferentes cometidos que unas y otras poseen, nos percataremos de que ha de existir un tajo entre ellos. Si han de cumplir sus respectivos propósitos, los asertos valorativos y los descriptivos han de ser diferentes, pues si aquéllos fueran objetivos ya no podrían fungir como valorativos. Dicho metafísicamente, los valores no pueden estar en el mundo, pues si lo estuvieran dejarían de ser valores y serían otra parte del mundo. Dicho de manera formal, no se puede definir una palabra valorativa en términos de las descriptivas, pues si ello fuera posible no se podría emplear ya la palabra valorativa para comendar, sino sólo para describir. Dicho todavía de otro modo, todo intento de derivar un 'debe' de un 'es' ha de ser pérdida de tiempo, pues todo lo que podría mostrar, puesto caso que lo lograra, sería que el 'es' no era un verdadero 'es', sino un 'debe' disfrazado o, en todo caso, que el 'debe' no era un verdadero 'debe', sino un 'es' solapado.

Este resumen del punto de vista tradicional ha sido muy sucinto, pero creo que trasunta algo del poder del cuadro. En manos de ciertos autores modernos, especialmente de Hare y Nowell-Smith, ese cuadro alcanza notable sutileza y elevación conceptual.

¿Qué hay de mal en el cuadro? No hay duda de que muchas cosas. Acabaré diciendo que una de las cosas que están mal es que no logra darnos razón coherente de nociones tales como compromiso, responsabilidad y obligación.

Para llegar a esa conclusión empezaré diciendo que ese cuadro no da razón de los *diferentes tipos* de proposiciones descriptivas. Sus ejemplos de proposiciones descriptivas son 'mi coche va a 60 km/h', 'Juan mide 1,70', 'Pedro tiene el pelo castaño', etc. Pero es forzado, por su propia rigidez, construir 'Juan se casó', 'Pedro hizo una promesa', 'Gómez tiene cinco pesos' y 'Alfonso metió un gol', como proposiciones descriptivas. Y es forzado porque que uno se case, haga una promesa o deje de hacerla, tenga o no tenga cinco pesos, meta el gol o no lo meta, es un hecho objetivo, como el que se tenga pelo rojo o castaño. Con todo, el primer tipo de proposición (las que contienen 'casarse', 'promesa', etc.) parecen ser muy diferentes de los paradigmas empíricos simples de las proposiciones descriptivas. ¿En qué se diferencian? Si bien los dos tipos de proposiciones plantean cuestiones de hechos objetivos, las proposiciones en que entran palabras como 'casarse', 'prometer', 'gol' y 'cinco pesos' hablan de hechos cuya existencia presupone ciertas instituciones: alguien tiene cinco pesos, puesto que está la institución de la moneda; haz desaparecer la institución y todo lo que le quedará será un rectángulo de papel con tinta de determinado color. Alguien mete gol, dada la institución del fútbol; hágase desaparecer éste y todo lo que hará será dar un puntapié a un balón. Similarmente, alguien se casa o promete dentro de las instituciones del matrimonio y de la promisión; sin ellas, todo lo que haría sería proferir palabras o hacer ademanes.

Podemos caracterizar tales hechos como hechos institucionales y contrastarlos con los no institucionales o hechos brutos: que alguien tenga un pedazo de papel con tinta de determinado color es un hecho bruto; que tenga cinco pesos es un hecho institucional[6]. El cuadro clásico no da razón de las diferencias entre las proposiciones de hechos brutos y las de hechos constitucionales.

Aquí, la palabra 'institución' suena a artificial, por lo que preguntaremos: ¿qué clase de instituciones son? Para poder responder a esta pregunta he de distinguir entre dos tipos de reglas o convenciones. Hay unas reglas que regulan formas de comportamiento previamente existentes. Por ejemplo, las reglas de etiqueta regulan la manera de comer, pero el comer existe independientemente de esas normas. Otras reglas, por el contrario, no regulan meramente, sino que crean o definen nuevas formas de comportamiento: por ejemplo, las reglas del ajedrez no regulan sólo una actividad anteriormente existente llamada jugar al ejedrez; por así decir, crean su posibilidad o definen esa actividad. La actividad de jugar al ajedrez se constituye por la acción que se amolda a esas normas; el ajedrez no tiene existencia independientemente de esas reglas. La distinción que hago fue prefigurada por la distinción que Kant hizo entre principios regulativos y constitutivos; así pues, adoptaremos su terminología y describiremos nuestra distinción como distinción entre reglas regulativas y constitutivas. Las reglas regulativas rigen actividades cuya existencia es independiente de las reglas; las reglas constitutivas constituyen (y también rigen) formas de actividad cuya existencia depende lógicamente de esas reglas[7].

Ahora bien, las instituciones de que he hablado son sistemas de reglas constitutivas. Las instituciones del

6 Para una discusión de esta distinción, ver G. E. Anscombe 'Brute Facts', *Analysis* (1958).

7 Para una discusión de una distinción conexa, ver J. Rawls, 'Two Concepts of Rules', *Philosophical Review*, LXIV (1955).

matrimonio, de la moneda y de la promisión se parecen a las del fútbol o del ajedrez en que son sistemas de tales reglas constitutivas o convenciones. Lo que he llamado hechos institucionales son hechos que presuponen tales instituciones.

Una vez que percibimos su existencia y empezamos a captar la naturaleza de tales hechos institucionales, sólo nos queda un breve paso para ver que existen muchas formas de obligaciones, de compromisos, de derechos y de responsabilidades, que están institucionalizadas de modo semejante. Es un hecho que se tienen ciertas obligaciones, compromisos, derechos y responsabilidades, pero se trata de hechos institucionales, no brutos. Y fue una de tales formas institucionalizadas de obligación, el prometer, la que yo invoqué arriba para derivar un 'debe' de un 'es'. Comencé con un hecho bruto, que un hombre profería ciertas palabras, y luego invoqué la institución de tal manera que generara hechos institucionales por los que llegamos al hecho institucional de que un hombre debía pagar a otro cinco dólares. Toda la prueba reposa sobre la apelación a la regla constitutiva que establece que hacer una promesa es asumir una obligación.

Estamos ahora en posición de ver cómo podemos generar un número indefinido de tales pruebas. Considérese el ejemplo siguiente tan diverso. Estamos en nuestra mitad de la séptima entrada y debo pasar a la tercera base. El *pitcher* lanza y la pelota va a dar al *shorstop*, y me tocan cuando ya he corrido unos tres metros. El *umpire* grita: '¡Fuera!', pero yo, como soy positivista, sigo en el campo. El *umpire* me ordena que me retire, pero le digo que no se puede derivar un 'debe' de un 'es'. Ningún número de proposiciones descriptivas de cuestiones de hecho —digo— implicará proposiciones valorativas respecto de que yo haya o deba dejar el campo. No puedes extraer órdenes o recomendaciones de hechos solos; lo que se precisa es una premisa mayor valorativa. Por tanto regreso y me quedo en la segunda base (hasta que me saquen

a rastras del campo). Creo que todos han de ver que mis alegatos son descabellados en este caso, y lógicamente absurdos. Claro que puedes derivar un 'debe' de un 'es', y aunque ponernos realmente a derivar la secuencia sería aquí mucho más complicado que en el caso de la promisión, en principio no hay diferencia alguna. Al comprometerme a jugar baseball me he sometido a la observación de ciertas reglas constitutivas.

Estamos ahora en posición de ver que la tautología de que se deben cumplir las promesas es sólo parte de una clase de tautologías similares concernientes a las formas institucionalizadas de obligación. Por ejemplo, 'no se debe robar' se puede tomar como afirmación de que reconocer algo como propiedad de alguien supone necesariamente reconocer su derecho a hacer de ello lo que quiera. Es una regla constitutiva de la institución de la propiedad. 'No se deben decir mentiras' puede tomarse como afirmación de que el hacer una aserción comporta necesariamente asumir la obligación de hablar verazmente. Es otra regla constitutiva. 'Se deben pagar las deudas' se puede construir como afirmación de que reconocer algo como deuda es reconocer necesariamente que existe la obligación de pagar. Es fácil ver cómo todos

8 Proudhon dijo: 'La propiedad es un robo'. Si se toma esto como un prenotando interno, carece de sentido. Fue propuesto como prenotando externo que atacaba y rechazaba la institución de la propiedad privada. Gasta aires de paradoja y obtiene su fuerza porque emplea términos que son internos de la institución con el fin de atacarla.

Desde la cubierta de unas instituciones se pueden hacer chapuzas echando mano de reglas constitutivas o incluso arrojar por la borda otras instituciones. Pero, ¿cabe echar por la borda todas las instituciones (con el fin, quizá, de no tener que derivar jamás 'debe' de 'es')? No sería posible y seguir aceptando aquellas formas de conducta que consideramos característicamente humanas. Supongamos que Proudhon hubiera añadido (e intentado vivir de conformidad con ello): 'La verdad es una mentira; el matrimonio es infidelidad; el lenguaje no comunica; la ley es crimen', y así sucesivamente de cada institución posible.

estos principios generarán contra-ejemplos de la tesis de que no se puede derivar un 'debe' de un 'es'.

Así, pues, mis conclusiones tentativas son:

1. El cuadro clásico no da razón de los hechos institucionales.
2. Los hechos institucionales existen dentro de sistemas de reglas constitutivas.
3. Algunos sistemas de reglas constitutivas comportan obligaciones, compromisos y responsabilidades.
4. Dentro de esos sistemas es posible derivar 'debe' de 'es', según el modelo de la primera derivación.

Con estas conclusiones podemos volver a la pregunta con que empecé esta sección: ¿Cómo concediendo un simple hecho respecto de alguien, como el hecho de que haya proferido ciertas palabras o que haya prometido, me puedo ver obligado a admitir que él deba hacer algo? Se puede empezar a responder a esta pregunta diciendo que establecer tal hecho institucional es invocar las reglas constitutivas de la institución. Son esas reglas las que dan a la palabra 'promesa' su significado. Pero esas reglas son tales que aceptar que Ticio hizo una promesa supone obligarme a aceptar que debe hacer algo (siendo iguales las demás cosas).

Si se quiere, pues. hemos mostrado que 'promesa' es vocablo valorativo, pero puesto que también es puramente descriptivo, hemos mostrado realmente que es preciso reexaminar toda la distinción. La supuesta distinción entre proposiciones descriptivas y valorativas es en realidad una fusión de al menos dos distinciones. Por una parte está la distinción entre las diferentes clases de actos del habla, siendo una familia de actos locutorios las valoraciones y otra las descripciones. Es una distinción entre diferentes clases de fuerza ilocucional [9]. Por otra parte, está la dis-

[9] Ver J. L. Austin, *How to Do Things With Words* (Cambridge, Massachusetts, 1962) donde se explica esta noción.

tinción entre expresiones que implican asertos sobre cuya verdad o falsedad se puede decidir objetivamente, y expresiones de asertos sobre los que no es posible decidir nada objetivamente, sino que son 'asuntos de decisión personal' o 'asuntos de opinión'. Se ha supuesto que la primera distinción es (debe ser) un caso especial de la segunda, que si algo tiene la fuerza ilocucional de una valoración, no puede ir implicado en premisas fácticas. Parte del propósito de mi tesis es demostrar que tal doctrina es falsa, pues las premisas fácticas puede implicar conclusiones valorativas. Si estoy en lo cierto, entonces la supuesta distinción entre expresiones descriptivas y valorativas, es útil sólo como distinción entre dos tipos de fuerza ilocucional, la de describir y la de valorar, y ni aun en ese caso es muy provechosa, pues si hemos de emplear esos términos estrictamente, se trata de dos clases de fuerza ilocucional entre centenares de clases de esa fuerza, y las expresiones de asertos de la forma (5) —'Ticio debe pagar a Smith cinco dólares'— no entraría de manera característica en ninguna de esas clases.

VIII

EL JUEGO DEL PROMETER

R. H. Hare

De *Revue Internationale de Philosophie*, No 70 (1964), pp. 398-412. Reimpreso con permiso del autor y de *Revue Internationale de Philosophie*.

Una de las cuestiones de más fundamento en torno a los juicios morales es si ellos, al igual que otros juicios de valor, se pueden derivar lógicamente de proposiciones sobre hechos empíricos. Como ocurre con las cuestiones filosóficas más importantes, se ha llegado con ésta a un punto a partir del cual toda discusión ulterior se escindirá en fragmentos, en ejemplos, argumentos y contraargumentos particulares. Este artículo trata de ser una contribución a la controversia. En reciente escrito, 'Cómo derivar «debe» de «es»' [1], el profesor J. R. Searle prueba una empresa que muchos otros antes que él han procurado

1 *Philosophical Review*, 1964. Debo agradecer los conocimientos que me ha brindado un artículo inédito que amablemente me entregó el profesor A. G. N. Flew, así como algunas argumentaciones útiles del propio profesor Searle. La argumentación de Searle, aunque no la puedo aceptar, es más plausible y da un tono moral más alto que la últimamente publicada por el señor MacIntyre, y que ha sido repetida en forma intrascendentemente variada por el profesor Black *(Phil. Rev.*, 1959 y 1964). Mientras Searle trata de demostrar lógicamente que hemos de cumplir las promesas, Black y MacIntyre quieren decirnos que debemos hacer todo lo que sea el único y solo medio de llevar a cabo *cualquier* cosa que podamos desear, o evitar *todo* lo que deseamos evitar.

llevar a cabo. Su argumentación, si bien me parece viciada, está expuesta con tal claridad y elegancia que retribuye con mucho el escrutinio que se le haga.

Nos propone a consideración la siguiente serie de proposiciones:

(1) Ticio profirió las palabras 'Con este te prometo, Cayo, pagarte cinco dólares'.
(2) Ticio prometió pagar a Cayo cinco dólares.
(3) Ticio se puso bajo (asumió) la obligación de pagar a Cayo cinco dólares.
(4) Ticio está bajo la obligación de pagar cinco dólares a Cayo.
(5) Ticio debe pagar cinco dólares a Cayo.

Nos dice respecto de esta lista que 'la relación entre una proposición y la siguiente, si bien no es en todos los casos de 'implicación', sin embargo no es sólo una relación contingente, y que la proposición adicional necesaria para convertir la relación en la de implicación no es preciso que contenga proposiciones valorativas, principios morales o cosas por el estilo' (p. 44) [2].

Aunque en su argumentación pueda haber otros pasos cuestionables, me concentraré en los que van de (1) a (2) y de (2) a (3). Una de las 'proposiciones adicionales' que intercala Searle entre (1) y (2) es

(1a) Bajo ciertas condiciones *C*, quienquiera que profiera las palabras (declaratoria) 'Con este te prometo pagarte, Cayo, cinco dólares', promete pagar a Cayo cinco dólares.

Esta, nos dice, en conjunción con otra premisa,

(1b) Las condiciones *C* han lugar,

convierte el paso de (1) a (2) en una implicación (páginas 44 *s.*). A continuación inserta de manera semejante entre (2) y (3), para mostrar que ese paso es una implicación, lo que llama la premisa 'tautológica' [3].

2 pp. 146-47 de este volumen [E.]

3 Parece que es preferible 'analítico', pero emplearé el término de Searle.

(2a) Todas las promesas son actos de situarse bajo (asumir) la obligación de hacer la cosa prometida.

Esta premisa es 'tautológica' porque 'No será completo ningún análisis del concepto de prometer si no contiene la característica de que el promisor se sitúa bajo una obligación' (p. 45)[4].

Más tarde, Searle plantea lo que parece ser el mismo punto, pero hablando de lo que llama 'reglas constitutivas'. Existen algunas instituciones que no sólo se rigen, sino que se constituyen por las reglas que las regulan. Así, 'las reglas del ajedrez, por ejemplo, no regulan sólo una actividad anteriormente existente llamada jugar al ajedrez; por así decir, crean su posibilidad o definen esa actividad' (p. 55)[5]. Las reglas de ajedrez y del baseball son ejemplos de reglas constitutivas, y también lo es 'la regla constitutiva de que hacer una promesa es asumir una obligación' (p. 56)[6].

Consideraré las relaciones entre (1a) y (2a). Con el fin de esclarecerlas, acudiré a la analogía del 'baseball', que Searle nos brinda muy auxiliadoramente (página 56). Habla de un conjunto de condiciones empíricas tal que, si han lugar, el beisbolista está 'out' quedando en la obligación de dejar el campo. Llamaré a esas condiciones '*E*', con el fin de esconder mi ignorancia de las reglas del baseball en que esas condiciones están especificadas. Lo que, en el caso del 'prometer', corresponde a las condiciones *E* del baseball, son las condiciones *C junto* con la condición de que la persona en cuestión ha de haber proferido las palabras 'Prometo, etc.'. Enumeremos las proposiciones del caso del 'baseball' de manera que correspondan con la numeración de Searle en el caso del 'prometer', distinguiéndolas de éstas por el ápice de 'pri-

4 [p. 148 E.]
5 [p. 160 E.]
6 [p. 161 E.]

ma'. Habrá, pues, una regla constitutiva del baseball a tenor de la cual

(1a') Siempre que un jugador satisfaga las condiciones *E*, queda 'fuera'. Y, puesto que no habrá análisis completo del concepto de *'fuera'* si no incluye la característica de que el jugador que está 'fuera' queda obligado a dejar el campo, podemos añadir la premisa 'tautológica',

(2a') Todos los jugadores que están fuera quedan obligados a dejar el campo.

Podemos simplificar el argumento combinando (1a') y (2a') en una regla constitutiva única,

(1a'*) Siempre que un jugador satisfaga las condiciones *E*, está obligado a dejar el campo.

Pues si se aplica directamente a (1a') la definición en virtud de la cual (2a') es una tautología, se convierte en (1a'*). Y, similaremente, en el caso del 'prometer', se simplificará la argumentación si combinamos (1a) y (2a) de manera que se forme una regla constitutiva única,

(1a») Bajo ciertas condiciones *C*, quienquiera que profiera las palabras (declaratoria) 'Con esto te prometo pagarte, Cayo, cinco dólares', se sitúa bajo (asume) la obligación de pagar a Cayo cinco dólares.

Esta regla podría asentarse de una forma general, omitiendo la referencia a Cayo; pero no hace falta que nos preocupemos por eso.

¿Cuál, pues, es el status de (1a*)? Cinco son las respuestas que merecen examen:

(a) Es una tautología:
(b) Es una proposición empírica y sintética respecto de lenguaje común;
(c) Es una prescripción sintética sobre el lenguaje común;

(d) Es una proposición empírica y sintética acerca de algo más que no es el lenguaje común;
(e) Es, o contiene implícitamente, una valoración sintética o prescripción que no versa sólo sobre el lenguaje común.

Seguramente, Searle mantendría (b). Yo defenderé (e). Como los argumentos que aplicaré contra (a), (b) y (c) son los mismos, no será preciso darlos por separado para sendas respuestas. Será necesario rebatir separadamente (d), pero no nos llevará mucho.

Empecemos analizando el status de la proposición análoga (1a'*). ¿Es una tautología? Existe ciertamente una tautología con la que se puede confundir con facilidad, a saber:

(1a'*+) *En (esto es, según las reglas del) baseball,* siempre que un jugador satisface las condiciones *E*, está obligado a dejar el campo.

Es una tautología, porque la definición de 'baseball' ha de decir más o menos 'es un juego con las siguientes reglas...' y a continuación una lista de normas, entre las que estará (1a'*) u otra equivalente. Pero esto no convierte a (1a'*), en que se ha omitido la parte en cursiva, en tautología. (1a'*) es un resumen de las reglas del baseball, y aunque puede ser que algunas de las reglas de un juego sean tautologías, es imposible que todas lo sean. Puesto que si así fuere, lo que tendríamos no serían las reglas para jugar un juego, sino las reglas (o, más estrictamente, ejemplificaciones de reglas) para hablar correctamente sobre el juego. Para conformarse a las reglas de un juego es necesario actuar, no meramente hablar, de cierta guisa; por lo tanto, las reglas no son tautológías.

Por la misma razón, como veremos, las reglas del baseball (y en particular (1a') y (1a'*) no pueden ser tratadas como proposiciones sintéticas, ni siquiera como prescripciones sintéticas, sobre el uso de palabras. Versan sobre cómo se juega o se debe jugar un juego.

Apliquemos ahora todo esto al caso del 'prometer'. Por paridad de raciocinio se ve claro que (1a*) no es una tautología, aunque es fácil confundirla con otra proposición (1a*+), que *es* una tautología. (1a*+) constará de (1a*), precedida por las palabras 'En la institución de la promisión' (podríamos decir, si no se prestara a malas interpretaciones: 'En el juego del prometer'). Esto es una tautología, porque no se puede extender a 'Según las reglas de una institución, cuyas reglas dicen «Bajo las condiciones *C*, quienquiera que profiera las palabras... (etc., como en (1a*))», bajo las condiciones *C*, quienquiera que profiera las palabras... (etc., como en (1a*))'. Pero (1a*) en sí no es tautología. Como antes, las reglas constitutivas de una institución pueden contener algunas tautologías, pero no todas pueden ser tautologías, si han de prescribir que la gente *actúe* de cierta manera y no de otra. Y, como antes, no debemos descarriarnos pensando que, pues es una tautología que el prometer sea una institución de la que (1a*) es regla constitutiva, (1a*) en sí es una tautología.

Como antes, y por razones análogas, (1a*) no es ni una proposición sintética ni una prescripción sintética sobre cómo se habla o se debe hablar. Precisamente porque tiene las consecuencias que Searle le adscribe, es más que eso.

Hay una disparidad aparente entre los casos de 'prometer' y del baseball que puede ser fuente de confusión. En el caso del baseball, la palabra 'baseball' no ocurre en (1a'*) y, por tanto, aunque (1a'*) en cierto sentido es exclusivamente de 'baseball', no es por lo mismo tautológica. Pero en el caso del 'prometer', (1a*) contiene la palabra 'promesa', y ello hace que sea más aceptable decir que (1a*), al ser en cierto sentido explicativa de la noción de prometer, es una tautología. Mayor es todavía esta aceptabilidad en el caso de (1a). La respuesta a esta objeción puede ayudar a esclarecer todo el procedimiento de introducir en el lenguaje una palabra del tipo de 'prometer'. La palabra se introduce por medio de una proposición

como (1a*). Pero no nos debemos llevar al error de pensar que esto convierta (1a*) en una tautología o en una mera proposición respecto del empleo de palabras. Pues, como veremos, es característico de palabras como 'prometer', que sólo tienen significado dentro de instituciones, que únicamente puedan introducirse en el lenguaje cuando se asiente a ciertas proposiciones sintéticas sobre cómo hemos de *actuar*. (1a*) es proposición de ese tipo. La palabra 'prometer', para poseer significado, depende de la proposición, pero la proposición no es verdadera solamente en virtud del significado de 'promesa'. De manera similar, la palabra 'fuera' depende, en lo que hace a su significado, de las reglas del baseball o del cricket, pero esas reglas no son tautologías en virtud del significado de 'fuera' y otras palabras por el estilo.

Sin embargo, puede parecer que con esto no llegamos a la raíz de la objeción, pues el argumento de Searle podría expresarse sin mencionar para nada la palabra 'prometer'. Podría sin más sustituir en (1a) la palabra 'prometer' por 'cargar con la obligación'. La aseveración se convertiría entonces en

> Bajo ciertas condiciones *C*, quienquiera que profiera las palabras (declaratoria) 'Con esto cargo con la obligación de pagarte, Cayo, cinco dólares', carga con la obligación de pagar a Cayo cinco dólares.

Sin duda, se podría decir que es innegable que se trata de una tautología o, si se quiere, de una aclaración respecto del empleo de palabras. Pero esto es precisamente lo que quiero negar. Pues, si en primer lugar, la mera repetición de las palabras 'carga... obligación' del aserto la convirtiera en tautológica, sería difícil comprender qué hacen las palabras 'Bajo ciertas condiciones *C*'. Se podría pensar que bajo cualesquiera condiciones, la persona que dijera 'Con esto cargo con la obligación de etc.', por lo mismo habría cargado con la obligación de etc. Pero una vez que vemos que esto no es así (por ejemplo la

persona puede estar bajo coacción o ser demente), comprendemos que la apariencia de tautología es engañosa. No es cierto en general (tanto menos, tautológico) que quien dice '*p*' hace realidad '*p*'. Cosa parecida sucede ahora con el verbo 'prometer'. Quien dice 'prometo', promete (bajo ciertas condiciones). Pero no es tautología que lo haga, ni es tautología que el individuo que dice 'Con este cargo con la obligación de' carga con una obligación, incluso bajo ciertas condiciones (empíricas). Tampoco son apostillas sobre el empleo de palabras, pues es necesario, para la adopción de estas expresiones de ejecución, que se adopten también ciertas reglas sintéticas constitutivas (y no meramente lingüísticas), creando así la institución dentro de la cual tienen significado esas expresiones.

Para dejar esto más claro, supongamos que poseemos en nuestro lenguaje la palabra 'obligación' (y palabras emparentadas, como 'debe'), pero que ninguna de nuestras obligaciones ha sido 'institucionalizada', como dice Searle (p. 56)[7]. Es decir, podemos hablar de tener obligaciones (*v. g.*, de alimentar a nuestros hijos) e incluso de cargar con obligaciones (*v. g.*, al tener hijos, cargamos con la obligación de alimentarlos); ahora bien, no podemos hablar aún de cargar con una obligación diciendo meramente 'Cargo con la obligación, etc.'. Supongamos que una persona ingeniosa insinúa la adopción de esta expresión útil (o más bien su conversión a este nuevo empleo). Los otros miembros de la sociedad se le pueden quedar mirando y decir 'Pero no vemos cómo puedes cargar con una obligación sólo diciendo esas palabras'. Lo que deberá decir para venderles este producto, y con él la institución de que es parte, será algo así: 'Habéis de adoptar la regla constitutiva o principio moral de que uno tiene obligación de hacer aquellas cosas sobre las que se ha dicho «(Con esto) cargo con la obligación de hacerlas».' Cuando hayan adoptado

7 [p. 160 de este volumen. E.]

este principio, o al adoptarlo, pueden introducir el nuevo empleo de la expresión. Y el principio será sintético. Es un principio moral sintético nuevo y no meramente un nuevo modo de hablar lo que se está introduciendo. Esto se manifiesta por el hecho de que, si adoptan el principio, habrán adquirido obligaciones de hacer cosas que no han realizado antes, no meramente de hablar de guisa que no han hablado antes.

Puede haber, ciertamente, una interpretación según la cual (1a), (1a*) y sus análogos podría decirse que son proposiciones 'sobre' el idioma. Podrían tratarse como proposiciones que dijeran o implicaran que en el idioma se tiene la expresión de ejecución 'Prometo', o la expresión de ejecución 'Me pongo bajo la obligación de', cuyo empleo está ligado a la institución de la promisión (o de asumir obligaciones), lo que por tanto supone que quienes hablan el idioma (o los que son idóneos) se suscriben a las reglas de la institución. La última mitad de ésta sería una declaración antropológica sobre los que hablan el idioma. Pero es obvio que tal proposición no podría generar implicaciones como las que Searle exige, pues las conclusiones que entonces se seguirían habrían de ser, a lo más, del tipo: 'Los que hablan el idioma se suscriben a la opinión de que Ticio está bajo obligación'; 'Los que hablan el idioma se suscriben al punto de vista de que Ticio debe', etc. Para que se sigan las conclusiones requeridas, no antropológicas, morales (o al menos prescriptivas), (1a) se ha de tomar —interpretada a la luz de (2a)— como que expresa la propia suscripción del hablante a las reglas de la institución del prometer, es decir, a los principios morales. No quiero discutir cuál es la manera más natural de tomar estas proposiciones; todo lo que tengo que decir es que *a menos que* se tomen de esta manera no funcionará la derivación.

Sucede con frecuencia que las expresiones de ejecución no pueden aplicarse sin la adopción de reglas constitutivas sintéticas. Así, sería imposible aplicar

la expresión 'Reclamo esta tierra', a menos que se adopte al mismo tiempo un principio de que, al decir eso, bajo las condiciones apropiadas, si el reclamante no ha sido anticipado por alguien más, adquiere al menos algún derecho sobre la tierra. En los días de los pioneros, en América se practicaba esto; ¡pero qué se hubiera hecho en Siberia, donde no regía ese principio!

Otra manera de mostrar que (1a*) no es una tautología, y se convierte en tal por el hecho de que se emplee para introducir en el lenguaje la palabra 'prometer', es la que sigue. Si (1a*) fuera verdadera en virtud del significado de la palabra 'prometer' y, por tanto, fuera tautológica, entonces tanto (1a) como (2a) tendrían que ser tautológicas. Pues se llegó a (1a*) aplicando a (1a) la definición que hizo tautológica a (2a), y es imposible extraer una tautología de una proposición sintética por sustitución definitoria. Pero (1a) y (2a) no pueden convertirse en tautológicas sin equivocar con la palabra 'prometer'. Pues (2a) es tautológica, si lo es, en virtud de *una* definición de 'prometer', y (1a) es tautológica, si lo es, en virtud de *una* definición de 'prometer' (o, en la otra suposición de que (1a) es una proposición sobre lenguaje, sólo puede serlo en virtud de *otra* definición de 'prometer'). Si tomamos (1a) como tautología, o como una proposición de uso, lo será en virtud de alguna definición como la siguiente:

(D_1) Prometer es decir, bajo ciertas condiciones, *C*, 'Con esto te prometo, etc.'.

Pero si (2a) es tautológica, lo es en virtud de una definición *diferente*, a saber,

Prometer es colocarse bajo una obligación... Cómo se completa la definición no tiene importancia; de todas maneras tiene que empezar así. Para convertir (1a*) en tautológica o en proposición de uso, tenemos que tomar simultáneamente 'prometer' en estos dos sentidos diferentes. Y no salimos del apuro completando así la última definición:

(D_2) Prometer es colocarse bajo obligación diciendo, bajo determinadas condiciones *C*, 'Con esto te prometo, etc.'.

Esta definición parece atractiva, y puede ser más o menos correcta, pero no convierte en tautología a (1a) y la convertiría en algo más que una proposición sobre uso de palabras. Según (D_2), quien diga 'Con esto te prometo, etc.', satisface sólo una de las condiciones de la promisión, pero a lo mejor no ha satisfecho la otra; puede haber dicho las palabras, pero no por eso ha tenido que cargar con obligación alguna. Sólo podemos decir que ha logrado esto, si asentimos al principio *sintético* (1*).

La necesidad de asentir a ese principio sintético, para que el dispositivo funcione, se puede solapar tomando (D_2) no como una definición verbal de tipo moderno, sino como el viejo artilugio de los aprioristas sintéticos, como una definición 'esencial' o 'real' de prometer. Pero entonces será sintética.

Concluyo por estas razones que (1*) no puede ser tautológica o una proposición sobre el empleo de las palabras, sino una regla constitutiva sintética de la institución de la promisión. Si las reglas constitutivas de la institución de la promisión son principios morales, como creo que lo son, entonces (1a*) es un principio moral sintético. Se sigue que, si Searle continúa sosteniendo que (2a) es tautológico, tiene que conceder que (1a) es o contiene implícitamente un principio moral sintético. Pero esto destruiría su tesis, y en efecto dice que no lo es, pues después de haberlo introducido dice 'hasta donde me es dado ver, no amaga ninguna premisa moral en este rimero lógico' (p. 43)[8]. Dice esto a pesar del hecho de que inmediatamente va a hacer (1a), por definición, equivalente a (1a*), que hemos visto es un principio moral sintético.

[8] [p. 148 de este volumen. E.]

Se puede insinuar que (1a) es una proposición empírica de alguna clase no-lingüística. Searle me aseguró que no cree que lo sea, pero este prenotando es digno de atención. Si fuera cierto, podría salvar su argumentación que, esencialmente, consiste en que no se ha de incluir ninguna premisa *no* moral o demás no empíricas y no tautológicas. Tiene algunos empeños en demostrar que las condiciones *C*, a que alude (1a), son empíricas, y esto se puede conceder como medio de argumentar. Pero si esto podría convertir la proposición (1b), 'Las condiciones *C* han lugar', en una proposición empírica, no opera lo mismo con (1a). Puesto que, por empíricas que esas condiciones *C* pudieran ser, es posible construir proposiciones no-empíricas, y aun incluso imperativos, de la forma 'Bajo las condiciones *C*, *p*'; por ejemplo, 'En condiciones *C*, desconectar (*o* se ha de desconectar) el motor'. No obstante, es fácil pensar erradamente que, si las condiciones bajo las cuales quien profiera 'Con esto prometo' se puede decir con razón que ha efectuado una promesa, son condiciones empíricas; lo que prueba que (1a) no es una aserción moral.

He dicho que concentraría mis ataques en los pasos del (1) al (3) del argumento de Searle. Pero diré de paso que se puede efectuar un ataque análogo contra los pasos del (3) al (5). También éstos dependen de una regla no tautológica de la institución del prometer o, en general, de colocarse (ejecutoriamente) bajo obligaciones. Esta regla no tautológica es como sigue:

(3a) Si alguien se ha situado bajo obligación (en el pasado), está (todavía) bajo obligación, a menos que haya efectuado lo que tenía obligación de hacer.

Para averiguar si esto es una tautología, tendríamos que reescribirlo, como antes, con auxilio de la definición o tautología que se requiere para convertir el paso de (4) a (5) en una implicación; a saber, la definición:

(D_3) Para que alguien esté bajo la obligación de hacer algo tiene que ocurrir el caso de que deba hacer eso.

(No escudriñaré si esta definición basta; probablemente no); o la tautología:

(4a) Todos los que están bajo la obligación de hacer algo, lo deben hacer.

(3a) Se convierte entonces en

(3a*) Si alguien se ha colocado bajo obligación (en el pasado), ocurre (todavía) que debe hacer aquello bajo cuya obligación de hacer está, a menos que ya lo haya efectuado. Que no es una tautología (o, por lo que nos toca, una proposición sobre empleo de palabras) se podría demostrar, si no está ya claro, mediante un argumento análogo al precedente.

Concluiré con algunas observaciones generales sobre la naturaleza de la equivocación que a mi parecer ha cometido Searle en su artículo. Hay muchas palabras que no podrían tener uso, a menos que los usuarios, o suficiente número de ellos, asienten ciertas proposiciones. La posibilidad del empleo de una palabra puede depender del asentimiento que se dé a proposiciones sintéticas. Esto tiene aplicación especialmente a muchas palabras cuyo empleo depende de la existencia de instituciones, aunque no sólo tiene aplicación con tales palabras [9]. Si no existieran leyes sobre la propiedad, no podríamos hablar de 'mío' y 'tuyo'; con todo, las leyes sobre la propiedad no son tautologías. A menos que exista aceptación de moneda a trueque de géneros, de nada servirían las palabras como 'dólar' y 'libra'; con todo, el aceptar moneda a cambio de géneros no es asentir a una tautología o a una aserción sobre lenguaje. En una comunidad que no jugara o no aceptara las reglas del baseball, la palabra 'fuera', como la usan los *umpires*, no tendría uso (aunque no como la emplearían antropólogos que hablaran de una comunidad que jugara baseball), pero esto no convierte las reglas del baseball

9 Quizá quepa insinuar que lo que Kant tenía en mente, sin el apriorismo sintético, era posiblemente que muchas palabras que empleamos en física y en la vida de cada día, como 'mesa' y, en general, 'objeto material', carecería de empleo a menos que hagamos ciertas presuposiciones sobre la regularidad del universo.

en tautologías o proposiciones sobre empleo del lenguaje.

En el caso de la promisión, tenemos fenómeno similar. A menos que un número suficiente de personas esté preparado a asentir a los principios morales, que son las reglas constitutivas de la institución de la promisión, la palabra 'prometer' carecería de uso. Supongamos el caso extremo de que nadie pensara que se deban cumplir las promesas. Entonces sería imposible hacer promesas. La palabra 'prometer' se reduciría a un simple ruido (excepto, como antes, en boca de los antropólogos), a menos que adquiriera nuevo uso. Pero de esto no se sigue que los principios morales, cuyo asentimiento por parte de suficiente número de personas es condición para el uso actual de la palabra 'prometer', no sean analíticos en sí.

Es necesario, además, que sólo un número suficientemente grande de individuos acepte la regla constitutiva. Si la acepta, y la palabra en cuestión entra en el uso, será posible que la gente que no se someta a esas reglas emplee la palabra con sentido. Así, un anarquista puede emplear la palabra 'propiedad'; un hombre que por razones muy propias no tenga confianza en el papel moneda y no quiera trocar mercancías con él, puede emplear la palabra 'libra', y un político maquiavélico que no reconozca la obligación de cumplir las promesas puede emplear la palabra 'prometer'. Incluso puede emplearla para hacer promesas, a buen recaudo de que no se conozcan sus opiniones morales.

Tales individuos son, como se sabe, parásitos, aunque no todos los parásitos son de reprender. Supongamos que alguien se oponga a la caza de zorras; esto no le impide que entre en la caza de zorras, en el sentido de acudir a las reuniones, seguir las partidas, etc., y emplear toda la terminología de la caza. Puede pensar que es su deber permitir, siempre que se tercie, que la zorra logre escapar (puede ser por esto por lo que participe en la caza); no obstante, esto no lo lleva a ninguna autocontradicción. Podrá ser

que ayudar a escapar a las zorras es contrario a las reglas constitutivas de la caza de zorras[10], pues a menos que entre esas normas haya una que diga que el objeto de la partida es matar a la zorra, no será *caza* de zorras. Pero esto no será óbice para que nuestro opositor de los deportes cruentos se presente como persona que acepta esas reglas, ni tampoco significa que al insinuarse de esa manera carga con obligación alguna de cumplir la regla. Y de la misma guisa, el político maquiavélico puede pensar sin autocontradicción que es su deber romper algunas de las promesas hechas (y pensarlo incluso mientras las profiere). No podría haberlas hecho, a menos que la palabra 'prometer' estuviera en uso, y no podría estar en uso si no existiera un número de personas que asistiera a los principios morales que gobiernan la promisión; pero esto no quiere decir que una persona que al hacer promesas disienta, tácitamente, de los principios, se contradiga. Al emplear la palabra 'prometer', por cierto, se disfraza de alguien que piensa que se han de guardar las promesas, de igual manera que quien miente se disfraza de alguien que piensa que *p* es, cuando no piensa así; pero ni el embustero ni el hombre que hace promesas mendaces se contradicen. Y cuando el promisor mendaz rompe su promesa, no se contradice con todo; puede decir 'Presumo pensar, cuando prometo, que se han de guardar las promesas, pero no lo pienso en verdad que las tenga que cumplir y nunca las cumplo'.

Hablar de 'hechos institucionales' si puede ser iluminador, puede ser también un modo insidioso de cometer la 'falacia naturalista'. No pienso que Searle caiga realmente en esta trampa en particular, pero sí que quizá caen otros. Existen principios morales, y otros principios, aceptados por la mayoría, que si no

10 Se puede objetar que las reglas de la caza de zorras no son constitutivas, sino regulativas. Esto dependerá de si estipulamos que existe alguna diferencia trascendente entre cazar zorras y cazar pelotas de cricket, cuestión en la que no entraré, pero cuya investigación puede ocasionar dudas respecto de esta distinción.

fueran aceptados de manera general no existirían algunas instituciones como la promisión o la propiedad. Y si existen las instituciones, estamos en posición de afirmar ciertos 'hechos institucionales' (por ejemplo, que cierta haza de tierra es mi propiedad), sobre la base de que ciertos 'hechos brutos' se cumplen (*v. g.*, que mis antepasados la han ocupado desde tiempos inmemoriales). Pero de los 'hechos institucionales' se pueden derivar ciertas conclusiones obviamente prescriptivas (por ejemplo, que nadie debe privarme de mi terreno). Así, parece como si hubiera una deducción directa en dos pasos, de los hechos brutos a las conclusiones prescriptivas, *vía* hechos institucionales. Pero esta deducción es un fraude. Pues el hecho bruto será motivo de conclusión prescriptiva sólo si el principio prescriptivo, regla constitutiva de la institución, es algo aceptado y ese principio prescriptivo no es tautológico. Pues si alguien (un comunista, por ejemplo) no acepta este principio descriptivo no tautológico, la deducción se desmorona como un castillo de naipes, aunque esto no le impida continuar empleando la palabra 'propiedad' (con ironía).

Lo mismo vale de la promisión. Puede parecer como si el 'hecho bruto' de que un persona ha proferido cierta secuencia fonética implica el 'hecho institucional' de que ha prometido y que esto, a su vez, implica que debe hacer ciertas cosas. Pero se puede deducir esto sólo si alguien acepta, además, el principio no tautológico de que se han de cumplir las promesas. Pues a menos que se acepte ese principio, no se es miembro subscripto de la institución que aquél constituye y, por ende, no puede ser compelido lógicamente a aceptar los hechos institucionales que genera, en sentido de que impliquen la conclusión, aunque es claro que se ha de admitir su verdad, considerados puramente como piezas de antropología.

Si no concuerdo con Searle en sus razones para mantener que hemos de cumplir las promesas, ¿cuáles son mis razones? Son de un carácter fundamental-

mente diferente, aunque de paso acepto partes de la argumentación de Searle. Romper una promesa es, de ordinario, una forma de decepción crasa. Es más crasa que el fracaso en cumplir una declaratoria de intención, precisamente porque (si así se desea) nuestra sociedad, *pari passu* con la introducción de la palabra 'prometer', ha adoptado el principio moral llamado 'promisión'. Mi razón para pensar que no he de tomar ventaja parasitaria de esta institución, sino que he de obedecer sus reglas, es la siguiente. Si me pregunto si aceptaría de grado ser engañado de igual guisa, sin dudar responderé que no. Por tanto, no me puedo suscribir a ningún principio moral que permita a la gente engañarse entre sí de esta manera (ningún principio general que diga 'Está correcto romper las promesas'). Tiene que haber principios más específicos que podría aceptar, de la forma 'Está correcto romper las promesas en situaciones del tipo *S*'. La mayoría acepta algunos principios específicos de esta forma. Cada uno ha de determinar qué sustituirá por '*S*', si sigue mi razonamiento, preguntándose —frente a dado valor de '*S*'— si puede suscribirse al principio cuando se aplica a todos los casos, incluyendo aquéllos en que ella es la persona a quien se hace la promesa. Así, la moralidad del cumplimiento de promesas es una aplicación bastante estándar de lo que en otro lugar [11] he denominado el tipo de 'regla de oro' de la argumentación moral; no es preciso que existan derivaciones 'es'-'debe' que la sostengan, derivaciones cuya validez solamente aceptarán aquellos que *a priori* hayan excluido cualquier interrogación acerca de las instituciones existentes sobre cuyas reglas se basan.

11 *Freedom and Reason*, esp. pp. 86-125.

IX

LA INTERPRETACION DE LA FILOSOFIA MORAL DE J. S. MILL[1]

J. O. Urmson

De *Philosophical Quarterly*, Vol 3 (1953), pp. 33-39. Reimpreso con la venia del autor y de *Philosophical Quarterly*.

Es asunto que debería interesar a quienes estudian la psicología de los filósofos, que las teorías de algunos grandes filósofos del pasado se estudien con la más paciente y acendrada erudición, mientras que las de otros son tomadas tan a la ligera y parodiadas, por parte de críticos y comentaristas, que es difícil creer que alguna vez se lean en serio, con interés de compenetración, o que siquiera se lean. Entre aquellos que más detrimento sufren por esta circunstancia es ejemplo conspicuo John Stuart Mill. Con excepción de un libro breve escrito por Reginald Jackson[2], no existe relación que remotamente sea cuidadosa de sus opiniones sobre la lógica deductiva, de manera que— por ejemplo— casi invariablemente se le hace padre de la opinión absurda de que el silogismo contiene *petitio principii*. Como dice Von Wright, 'Aún no se ha escrito una buena monografía sistemática y crítica de la lógica de la inducción de

1 [Este artículo se discute en el artículo de H. J. McCloskey 'An Examination of Restricted Utilitarism', *Philosophical Review* (1957) E.]

2 *An Examination of the Deductive Logic of J. S. Mill* (1941).

Mill'[3]. Pero todavía ocasiona mayor perplejidad la mala construcción, casi universal, asentada sobre las doctrinas éticas de Mill, pues su *Utilitarismo* es obra que ha de leer todo estudiante de filosofía, razón mayor para esperar que los críticos de Mill la hubieran leído siquiera una vez. Mas, por lo que se ve, no es así, y en vez de discutirse las doctrinas de Mill, se habla sobre un doble de las mismas, tanto que las críticas que se suelen hacer no tienen nada que ver con él. No será la tesis de este artículo sostener que las doctrinas de Mill estén inmunes de crítica, o que sean de clareza y consistencia verbal impecables. Sólo se propugnará que si se interpretaran con la mitad de la empatía que automáticamente se entabla con Platón, Leibniz y Kant, se descubriría una tesis esencialmente consistente, superior a la que se atribuye a Mill, y que quedaría inmune a las críticas comunes.

Otra advertencia se ha de hacer respecto de la finalidad de este artículo. Mill se propone dos cosas en su *Utilitarism;* en primer lugar, quiere dejar en claro el lugar de la concepción del *summum bonum;* en segundo lugar, intenta dar razón de la naturaleza del último fin. Sólo haremos de nuestra incumbencia la primera de estas dos partes de la teoría ética de Mill. No preguntaremos cuál era el fin último para Mill ni cómo pensaba que se pudiera establecer su punto de vista al respecto, sino qué parte, en su opinión, debía representar en una teoría ética sana la noción de fin último. Esta sección de la doctrina de Mill es independiente lógicamente de su disertación sobre la felicidad.

Dos interpretaciones equivocadas de Mill

Algunos de los expositores y críticos de Mill han pensado que éste intentaba analizar o definir la

3 *A Treatise on Induction and Probability* (1951), p. 164.

noción de correcto en términos del *summum bonum*. Así, Mill se presenta de ordinario como paradigma del naturalista ético cuando se interpreta naturalistamente su explicación de la felicidad, como si hubiera definido lo correcto atendiendo a las consecuencias naturales de las acciones. Moore, por ejemplo, al criticar las razones de Mill respecto del fin último, dice: 'Al insistir en que correcto ha de significar lo que produce los mejores resultados, se justifica plenamente el utilitarismo' [4]. Otros han sido menos favorables en el aprecio de esta supuesta opinión de Mill; pero, esté aceptada o no, me parece claro que Mill no la sostuvo. La única referencia de Mill a este problema analítico está en la página 27 (de la edición Everyman, a que aludirán todas las referencias), donde habla de una persona 'que viera en la obligación moral un hecho trascendente, una realidad objetiva perteneciente a la provincia de «las cosas en sí»', y sigue comentando esta manera de ver como carente de relación en absoluto con 'este punto de la Ontología' como si el análisis de los términos éticos no fuera parte de la filosofía ética cual la concebía, sino de la ontología. Parece claro que cuando Mill habla de que sus pesquisas versan sobre el 'criterio de correcto e incorrecto' (p. 1), 'respecto del fundamento de la moralidad (p. 1), con el fin de hallar 'una piedra de toque de lo correcto y equivocado' (p. 2), busca un 'medio de asegurarse qué está correcto y qué no lo está' (p. 2), no la definición de esos términos. No trataremos más de esta interpretación de Mill; si se requiere una refutación ulterior de ella, se habrá de buscar en la correspondencia del texto con la exposición distinta que en breve se dará.

El otro punto de vista equivocado evita el error de este primer punto de vista y, ciertamente, es incompatible con él. Es probablemente la opinión aceptada. Según esta interpretación, Mill busca una prueba de lo correcto e incorrecto como prueba última, por la

4 *Principia Ethica*, reimpreso en 1948, p. 106.

que se pueda justificar la adscripción de correcto o equivocado a las acciones, donde se supone que correcto y equivocado son palabras que entendemos. Con esa prueba se intenta ver si la acción tiende o no tiende a promover el fin último (que, sin duda, Mill dice que es la felicidad). Hasta aquí nada hay que objetar a esta mira aceptada, pues sin duda es atinada; pero en detalle está equivocada, pues se dice, además, que para Mill la última prueba es también la prueba inmediata, se ha de determinar lo correcto o errado de una acción en particular, considerando si secunda el fin último. Según Mill, hemos de admitir que a veces actuamos a ojo o apresuradamente, sin ponernos expresamente esta pregunta, pero la justificación real, si la hay, ha de ser directamente atendiendo a las consecuencias, incluidas las consecuencias del ejemplo que hemos puesto. De acuerdo con esto, Mill sostiene que una acción, una en particular, estará correcta si secunda el fin último mejor que cualquier otra, y si no es así, está equivocada. Por mucho que aderecemos nuestra mente en las situaciones morales, por lo que a la justificación se refiere, no entra en el asunto ningún otro factor. Es claro que esta interpretación de Mill queda abierta inmediatamente a dos objeciones que la desbaratan; en primer lugar, se apremia, como es natural y correcto, que si —*v. g.*— alguien ha efectuado una promesa, tiene que cumplirla no meramente por las consecuencias, incluso si esas consecuencias incluyen el ejemplo propio de romper la promesa. En segundo lugar, se señala con acierto que, según esto, el individuo que —*caeteris paribus*— escoja la inferior entre dos comedias musicales para una representación vespertina comete un mal moral, lo que es absurdo[5]. Si fuera ésta en efecto la opinión de Mill, valdría

5 Para un ejemplo de esta interpretación de Mill y de la primera y más importante objeción, ver Carritt, *The Theory of Moralts*, cap. iv.

poco más que para la erística renqueante de los niños sabihondos.

Interpretación corregida de Mill

Empezaré con una serie de proposiciones que, a mi manera de ver, son en efecto la doctrina de Mill, y las condensaré después, habida cuenta del contexto; esto obnubilará las sutilidades, pero esclarecerá los lineamientos principales de su interpretación.

A. Una acción particular se justifica como correcta si se demuestra que está de acuerdo con alguna regla moral. Se demuestra que está mal, señalando que transgrede alguna regla moral.

B. Se dice que una regla moral es correcta cuando se demuestra que reconocerla promueve el bien último.

C. Las reglas morales sólo se pueden justificar respecto de asuntos que rozan de manera considerable el bien común.

D. Donde no es aplicable ninguna regla moral, no tiene objeto suscitar la cuestión de la razón o error de los actos particulares, aunque se puede apreciar por otros medios cuál es el valor de las acciones.

Se ha de señalar como pormenor terminológico que cuando arriba aparece la frase 'regla moral', Mill emplea la expresión 'principio secundario' por lo general, aunque a veces dice también 'ley moral'. Con esos términos, de igual preferencia, Mill se refiere a preceptos como 'Guardar las promesas', 'No matar' o 'no decir mentiras'. En *On Liberty* (p. 135) se encontrará una lista de lo que Mill aprueba.

No hay duda de que es preciso explicar más estas proposiciones, pero se hará mejor, a la vez que se insinúan algunas cauciones, en el proceso de esclarecer los que de hecho son los puntos de vista de Mill. En primer lugar, pues, pasaremos a asentar que se deduce del texto de Mill que, a su manera de ver, las acciones particulares son correctas o equivocadas si

se puede demostrar que están de acuerdo o disceptan de alguna regla moral.

(i) Dice con evidente aquiescencia en la p. 2: 'La escuela intuitiva, lo mismo que la que se podría denominar inductiva, de ética insiste en la necesidad de leyes generales. Ambas están acordes en que la moralidad de una acción individual no es cuestión de percepción directa, sino de la aplicación de una ley a caso individual. Reconocen también en gran medida las mismas leyes morales'. Mill sólo echa en cara a estas escuelas que no logren dar un *rationale* unificador de esas leyes (como lo hará en la proposición B).

(ii) Dice en la p. 22: 'Pero una cosa es considerar las reglas de moralidad como improbables, y otra pasar por entero por sobre las generalidades intermedias, e intentar la prueba de cada acción individual directamente por el primer principio. Es una noción peregrina que el reconocimiento de un primer principio es inconsistente con la admisión de los secundarios'. Añade con sentimiento: 'La gente debería cesar de hablar sandeces a este respecto, pues ni las dirían ni las escucharían en otros asuntos prácticos'.

(iii) Habiendo admitido en la p. 23 que 'las reglas de conducta no pueden disponerse de tal manera que no admitan excepciones', añade (p. 24): 'Hemos de recordar que sólo en estos casos de conflicto entre los principios secundarios es ineludible apelar a los primeros principios. No hay caso de obligación moral en que no entre algún principio secundario; y si sólo entra uno, raramente habrá duda real sobre cuál es, si se trata de una persona que acepta dicho principio'. Esta cita va en apoyo tanto de la proposición A como de la D. Muestra que, para Mill, las reglas morales no son meramente cálculos de buen cubero que ayudan al hombre irreflexivo a arreglárselas, sino que son parte esencial del razonamiento moral. El hecho de que exista regla moral nos indica si estamos ante un caso de bien o mal, o ante otra situación moral o prudencial.

(iv) El último pasaje que elegiremos para determinar esta interpretación de Mill (sería fácil hallar más) es también una confirmación conjunta de las proposiciones A y D, donde se manifiesta que el último citado no fue un *obiter dictum* * sobre el que hubiéramos recargado demasiado peso. En el capítulo intitulado 'Sobre la conexión entre justicia y utilidad', Mill defiende que es indicio distintivo del acto justo el que sea requerido por una regla o ley específica, positiva o moral, que conlleva la sujeción a sanciones penales. A continuación escribe este importante párrafo (p. 45), que en vista de su momento y de la incuria que ha padecido citaremos por entero: 'Lo anterior es, según creo, razón verdadera, en lo que toca, del origen y crecimiento progresivo de la idea de justicia. Pero hemos de observar que hasta el momento no contiene nada que distinga esa obligación de la obligación moral en general. Pues es cierto que la idea de sanción penal, que es la esencia de la ley, no entra sólo en la concepción de injusticia, sino también en la de toda especie de error. No diremos que una cosa está equivocada, a menos que queramos dar a entender que alguien debe ser castigado de alguna manera por haberla efectuado; si no por la ley, por la opinión de sus prójimos; si no por la opinión, por los reproches de su propia conciencia. Este me parece que es el canto (*turning point*) real que distingue la moralidad de la conveniencia (*expediency*). Es parte de la noción de Deber en cada una de sus formas el que se pueda compeler a alguien a cumplirlo con todo derecho. El Deber es algo que se puede exigir de alguien, como se le exige que pague una deuda. Si no creemos que se le pueda exigir, no podremos decir que es un deber... Hay otras cosas, por el contrario, que nos gustaría que la gente hiciera, o que admiramos o nos place que sean hechas, o bien, nos disgusta o despreciamos a los demás si no las hacen, aunque confesemos que no tienen obligación

* Dicho de paso. (T.)

de hacer. Como no se trata de un caso de obligación moral, no los culpamos, es decir, no juzgamos que sean objetos apropiados de castigo... Creo que no hay duda de que esta distinción está a la base de las nociones de correcto y equivocado, pues llamamos a una conducta equivocada, o empleamos en su vez algún otro término de disgusto o de discrepancia, según creamos que la persona debe o no debe ser castigada por ello, y decimos que sería correcto actuar así y así o, meramente, que sería deseable o loable, según deseáramos ver a la persona, a la que le incumbe, compelida o sólo persuadida o exhortada, a actuar de tal manera'. Cómo los factores del punto de vista aceptado lo han hecho concordar con este pasaje es algo que no sé; tampoco lo mencionan. Si lo han advertido, presumiblemente lo han considerado como ejemplo del eclecticismo inconsistente de Mill. Dice bien claro Mill aquí que, a su modo de ver, lo correcto y lo equivocado se derivan de reglas morales. En otros casos en que queda afectado sin duda alguna el fin último, se ha de hacer el aprecio de la conducta por otros medios. Por ejemplo, si la participación de alguien queda menoscabada sin ruptura de la ley moral, se tratará (*Liberty*, p. 135) de imprudencia o de falta de respeto propio, pero no de acción mala. Baste esto como esclarecimiento de la interpretación de Mill, de manera positiva, por lo que respecta a los puntos A y D. Debemos preguntarnos ahora si hay algo en Mill que no esté de acuerdo con esto y que secunde el punto de vista aceptado.

Es imposible mostrar de manera positiva que no hay nada en Mill que favorezca el punto de vista aceptado, en contra de la interpretación dada aquí, pues exigiría revisión completa de todo lo que dice. Nos contentaremos con examinar dos puntos que podría pensarse apoyan el punto de vista aceptado.

(*a*) En la p. 6 dice: 'El credo que acepta como fundamento de la moral la Utilidad o el Principio de la Gran Felicidad, sostiene que las acciones son correctas en la proporción con que tienden a fomentar la

felicidad, y equivocadas si tienden a secundar lo contrario de la Felicidad'. Esta parece ser la bien conocida proposición que subyace en la interpretación aceptada. Por supuesto que se tomaría como una aserción laxa o imprecisa del punto de vista aceptado, si el argumento la requiriera. Pero adviértase que se puede decir estrictamente que cierta acción tiende a producir determinado resultado si se habla sólo de acciones-tipo y no de acciones-muestra. El beber alcohol suele producir jovialidad, pero el que beba este vaso la produce o no la produce. Parece, pues, que se puede interpretar aquí a Mill como considerando las reglas morales como tipos de acción que prohiben o son deleitosos; es decir, como señalando que reglas morales correctas son aquéllas que secundan el fin último (mi proposición B), sin decir algo contrario a la proposición A. Y esto, o algo como esto, es la interpretación que se requiere para que haya consistencia. La referencia de Mill a 'tendencias de acciones', al principio de la p. 22, refuerza el énfasis puesto aquí sobre la palabra 'tender', y ese contexto debería ser examinado por aquellos que exigen convicción ulterior.

(*b*) Mill a veces designa las reglas morales como 'generalizaciones intermedias' (*v. g.*, p. 22) del principio supremo, o como 'corolarios' del mismo (p. 22 también). Son éstas probablemente el tipo de frases que llevan a muchos a pensar que juegan un papel puramente heurístico en el pensamiento ético de Mill. Por lo que hace a la expresión 'generalización intermedia', no hay duda de que Mill piensa que deberíamos, y hasta cierto punto lo conseguimos, llegar y mejorar nuestras reglas morales por métodos como la observación de que cierto tipo de acción ha tenido malos resultados de carácter social en tal abrumadora mayoría de casos que se ha de descartar. (Pero esto es una simplificación fácil; véase la nota de la página 58 sobre cómo se ha de llegar a las reglas morales, y la relación pesimista sobre cómo llegamos de hecho a las mismas en *Liberty*, p. 69-70). Pero esta

disertación de la génesis de las reglas morales no exige que las interpretemos más que como reglas que fueron estatuidas alguna vez. Parece innecesario en realidad decir mucho de la expresión 'corolario'; obviamente no intentaba Mill que se tomara literalmente; de hecho es difícil determinar con exactitud cuál es la relación que existe entre las reglas morales y algún principio justificante, ni se esforzó Mill por hacerlo en un artículo popular en *Fraser*.

Las reglas morales y el fin último

En nuestro examen de las posibles objeciones a la proposición A ya hemos dicho algo en defensa del punto de vista por el cual, según Mill, una regla moral es correcta cuando su aceptación secunda el fin último (proposición B). Algo más puede decirse sobre esto, aunque parece bastante claro que si tenemos razón en decir que el principio supremo no debe ser evocado, según Mill, para justificar directamente actos correctos particulares, debe aparecer de manera indirecta, vista la importancia que Mill le daba. Es difícil pensar cuál puede ser la manera indirecta, si no es ésta. (i) En la p. 3, Mill reprocha a otros filósofos morales por no dar razón satisfactoria de las reglas morales, habida cuenta de un principio fundamental, aunque sitúan correctamente las reglas morales cual gobernadoras de las acciones particulares. Sería marchamo de filósofo inconsistente si no tratara de reparar la omisión seria que adscribe a los otros. (ii) Mill adjudica a Kant (p. 4) el empleo de argumentos utilitaristas, porque —afirma Mill— de hecho apoya las reglas de moralidad mostrando las malas consecuencias de no seguirlas o de seguir otras. Así, Mill considera aquí como claramente utilitarista la justificación o rechazo de las reglas morales atendiendo a sus consecuencias. No podría haber insinuado que Kant debió justificar directamente, aun sin sentirlo, las acciones particulares sobre tales mo-

tivos. Pero quzá no tenga propósito insistir más en este punto. Si alguien se ha convencido por lo dicho hasta aquí, no necesitará que se vuelva sobre lo mismo; con los demás será de más intentarlo.

¿A qué campos son aplicables las reglas de lo correcto y lo equivocado?

La aplicabilidad de las reglas morales, dice Mill, es la característica que diferencia no la justicia, sino la moralidad en general, de las restantes provincias de la Conveniencia y de la Recomendabilidad' (p. 46). Poco o nada dice en *Utilitarism* respecto de los límites entre moralidad y recomendabilidad (¿habría sido mejor, sin duda, haber dicho entre correcto e incorrecto, y los demás modos de aprecio moral y no-moral?). Parece razonable suponer que habría aceptado que el empleo de reglas morales debe confinarse a asuntos en que el tipo de consecuencia es lo suficientemente invariable para que no haya demasiadas excepciones. Pero ésta es una limitación pragmática; Mill tiene algo que decir acerca de una limitación en principio en *Liberty*, que he resumido cruelmente en mi proposición C (las reglas morales sólo se pueden mantener de manera justificada atendiendo a asuntos en que el bienestar general queda afectado más que desatendiblemente.

Es importante advertir que Mill en *On Liberty* habla de la libertad de sanciones morales, así como de las sanciones de la ley positiva. La distinción entre acciones auto-concernientes y las demás, a su entender, respecta tanto a la filosofía moral como a la política. El pasaje más notable que trata de la finalidad de las reglas morales está en la página 135. Aquí menciona cosas cual la intrusión en los derechos de los otros como 'objetos apropiados de repudio moral y, en casos graves, de retribución o de castigo moral'. Pero las faltas auto-concernientes (gustos bajos y demás) 'no son propiamente inmoralidades y

por fuertes que sean los tintes, no constituyen maldad... El término deber para consigo mismo, cuando significa algo más que prudencia, se refiere al autorespeto y auto-desarrollo'. Las faltas auto-concernientes convierten al culpable 'necesaria y propiamente en sujeto de disgusto o, en casos extremos, incluso de desprecio', pero esto pertenece a la esfera de la recomendabilidad, no de lo correcto o errado.

Baste esto sobre la disertación de Mill acerca de la lógica del razonamiento moral. He de recalcar que no se ha intentado otra cosa que una sinopsis de la respuesta de Mill, pues habla del asunto más rica y sutilmente en su libro. Respecto de la interpretación general, se ha de conceder más lugar a la lectura continuada, a la luz de esta sinopsis, que llevarla a cabo sobre la base de las pocas directrices que se han expuesto en este artículo. Es de afirmar categóricamente que no ha sido el propósito de este escrito propugnar que Mill ha finiquitado correctamente estos temas, quedando inmune a la crítica; ha sido sólo mi intención dar unas aclaraciones benévolas, sin hacer crítica ni en pro ni en contra. Pero sostengo sin duda alguna que las interpretaciones corrientes del *Utilitarism* de Mill son tan desaprensivas y van tan erradas, que la mayoría de las críticas que, de hecho, se basan en ellas carecen de valor y no lo rozan siquiera.

X

INTERPRETACIONES DEL 'UTILITARISMO' DE MILL

J. D. Mabbott

De *Philosophical Quarterly*, Vol. 6 (1956), pp. 115-20. Reimpreso con la venia del autor y de *Philosophical Quarterly*.

El artículo del profesor Urmson 'La interpretación de la filosofía moral de J. S. Mill' en *The Philosophical Quarterly* de enero de 1953 (Vol. 3, No. 10) es un trabajo sumamente interesante y estimulador. La tesis principal de Urmson es que los críticos anteriores han hecho sostener a Mill, como sin duda sostuvo G. E. Moore, que 'es siempre deber de todo agente llevar a cabo aquella acción, entre las que *pueda* efectuar en determinado momento, cuya *consecuencia total* posea el mayor valor intrínseco' (Moore, *Ethics*, p. 232). Pero, en opinión de Urmson, la posición auténtica de Mill era la siguiente:

'A. Una acción particular se justifica como correcta mostrando que está de acuerdo con alguna muestra moral. Se demuestra que está mal, señalando que transgrede alguna regla moral.'

'B. Se dice que una regla moral es correcta cuando se demuestra que reconocerla promueve el bien último (es decir, la mayor felicidad del mayor número)' (p. 35) [1].

1 [p. 186 de este volumen. E.]

Creo que hay que hacer dos ligeras enmiendas a la segunda cláusula. No basta con 'reconocer'; se requiere la práctica de acuerdo con esa ley. Y 'promover' sugiere que todas las reglas morales defendibles se han de reconocer u obedecer de hecho; insinúo 'promovería' (al menos como elucidación de Mill).

Ahora bien, de esos dos principios se siguen dos diferencias principales entre la interpretación ortodoxa del utilitarismo y la de Urmson. (1) Según la interpretación ortodoxa, nunca es correcto realizar una acción cuando hay otra optable que rendiría mayor bien (cf. la interpretación mooreana antes citada). Mas, según Urmson, puede ser correcto efectuar una acción que esté concorde con alguna regla moral, incluso si esa acción particular produce menos bien que otra acción optable, debido a que la práctica general de la regla produce más bien que la omisión de tal práctica o la práctica de otra regla optable. (2) Según la interpretación ortodoxa (comparar de nuevo con G. E. Moore), lo correcto de una acción se determina por sus consecuencias *reales;* según la interpretación de Urmson, en cambio, por sus consecuencias *hipotéticas,* por lo que *sucedería si* por lo general se practicara la regla que sigue la acción.

Hay un pasaje en *Utilitarism* (Everyman Edition —a que remitirán las demás citas—, pp. 17-18) en que Mill acepta explícitamente estos dos importantes corolarios. Aunque Urmson no lo trae, es uno de los fragmentos más notables en favor de su interpretación. 'En el caso de las abstinencias —de cosas que la gente prohíbe hacer por consideraciones morales, *aunque las consecuencias en el caso particular pueden ser beneficiosas*— sería indigno de un agente inteligente no percatarse de que la acción es de un tipo que, *si se practicara como regla general, sería nocivo por lo común,* y que éste es el motivo de la obligación de abstenerse de ello.

Si se vuelve a leer a Mill a la luz de los comenta-

rios de Urmson, quedan al descubierto muchos pasajes, como éste que va en su favor, cuya trascendencia parece haber escapado a los críticos anteriores. Pero me parece dudoso que la opinión de Mill sea clara y consistentemente la que Urmson propone. Hay muchos pasajes que cuadran en la interpretación ortodoxa antigua, y dudo de que el propio Mill advirtiera las diferencias fundamentales de una manera de ver y otra. Lo que resta de este artículo lleva el propósito no sólo de mostrar las dificultades que a la tesis de Urmson presentan algunos pasajes de Mill, sino también emplear esas dificultades para hacer resaltar más conspicuamente las diferencias entre las dos maneras de ver.

El punto principal de la nueva interpretación es que el primer principio no lleva a determinar lo correcto de algún acto particular. Mill dice que existe sólo una excepción al respecto; a saber, el caso en que dos reglas entran en conflicto. 'Hemos de recordar que sólo en estos casos de conflicto entre los principios secundarios es ineludible apelar a los primeros principios. No hay caso de obligación moral en que no entre algún principio secundario; y si sólo entra uno, raramente habrá duda real sobre cuál es' (p. 24). Pero cuando son dos las reglas que entran en conflicto, ¿qué he de preguntar? ¿Cómo aplicaré el primer principio para escaparme al dilema? ¿Me preguntaré si guardar una regla *en general* rinde más bien que cumplir la otra? Según la interpretación de Urmson, ésta parecería ser la pregunta correcta, pero sería muy difícil responderla. ¿O me preguntaré si cumplir una regla *en esa ocasión particular* me hará más bien que cumplir la otra? Pero de la misma manera podría haber hecho de lado toda referencia a las reglas y preguntar simplemente si A, que resulta concordar con la regla X, hará más bien que B, que resulta concordar con la regla Y. Mill no da indicio alguno sobre qué alternativa aprueba.

El pasaje arriba citado sostiene que la *única* excepción al ordenamiento de decidir acciones particu-

lares refiriéndonos al primer principio es cuando ocurre conflicto de principios secundarios. Pero existe otra excepción, que Mill permite en otro lugar. La 'excepción principal' a la regla contra el mentir se dice que es cuando callar la verdad 'salvará a un individuo de un mal grande y no merecido' (p. 21). La palabra 'no merecido' puede parecer que comporta un principio secundario conflictivo: 'a cada uno según se debe', pero no creo que esto tenga importancia. Mill admite que todos conocerán que cuando las consecuencias de cumplir una regla secundaria son muy malas (o romperlas es muy bueno), cabe la excepción. Ahora, esta otra excepción (que se denomina 'excepción principal') produce también una dificultad ulterior en la interpretación de Urmson. Dice además Mill, en el pasaje antes citado de la p. 24, que no existe caso alguno de obligación moral en que no entre algún principio secundario. ¿Qué decir del caso en que no entre ningún principio secundario y, no obstante, algún acto al que tengo acceso me pueda producir muy buenos resultados o aportar otros muy malos? ¿No será moral, correcto, deber mío, tal acto? Con todo, el único principio que entra aquí es el primer principio. Puede recordarse que junto a los deberes *prima facie* de la lealtad, etc., que corresponden a los principios secundarios de Mill, Sir David Ross alista los deberes *prima facie* de la beneficencia y de la no-maleficencia. Un modo de plantear estas dos dificultades es que, según la interpretación que de Mill hace Urmson, la producción de la mayor felicidad tendría que ser (*a*) una obligación *prima facie* (es decir, relacionada con la determinación de la corrección de actos particulares), (*b*) la base de cualquier otra obligación *prima facie* (o principio secundario), (*c*) el árbitro entre obligaciones *prima facie* conflictivas.

La tercera dificultad, admitida por Urmson, es que Mill llama 'corolarios' del primer principio a los secundarios (p. 22). Pero difícilmente serán corolarios si en un caso particular contradicen el primer prin-

cipio, cuando me abstengo de un acto particular con el fin de obedecer una regla, 'aunque las consecuencias del caso particular pueden ser beneficiosas' (página 18, arriba citada). El vocable 'corolario' indica, como concede Urmson, que el valor de los principios secundarios es puramente heurístico, y esto se infiere de las metáforas de Mill. 'Es una noción extraña que el reconocimiento de un primer principio sea inconsistente con la admisión de los secundarios... Indicar a un viajero su destino no es prohibir el uso de hitos y jalones a lo largo del camino' (pp. 22-3). Pero un hito o un jalón, en una ocasión particular, pueden no indicar el mejor camino a un lugar. Puedo andar a pie y existir un atajo por el monte, o la carretera con señales puede estar bloqueada por inundaciones o corrimientos de tierras. Diríamos entonces: 'no hagas caso del jalón'. Pero ¿qué ocurre cuando aplicamos la metáfora? El destino es la mayor felicidad para el mayor número; el jalón es la regla secundaria. ¿Qué sucede cuando una señal no indica el verdadero camino? ¿La hemos de preterir? Según la interpretación de Urmson, Mill diría: 'No, hay ocasiones en que, puesto el caso que haya otra regla conducente a la felicidad general, se ha de seguir la señal, la regla secundaria.' De manera similar con la comparación del almanaque (que ahorra al navegante calcular cada vez qué derrota seguir). No hay problema si se supone que el almanaque es infalible. Pero el almanaque de los principios secundarios no da directrices de marear que pongan rumbo a la mayor felicidad. Sin embargo, cuando no las da, Mill ha de sostener (según la interpretación de Urmson) que las hemos de seguir.

Se puede insinuar, para salvar la dificultad, como lo hacen Burke y G. E. Moore (*Principia Ethica*, página 162), que la razón de por qué habríamos de seguir una regla, incluso cuando romperla produciría mejores consecuencias a todas vistas, es que la regla guarda la sabiduría acumulada de generaciones, con sus experiencias y tradiciones, y que el individuo, por

tanto, tiene mayor probabilidad de errar en sus juicios, que de romperla obren mejores consecuencias, especialmente dado que su juicio puede estar afectado por sesgos o prejuicios. Pero es fácil dar con casos en que quedan excluidos tales sesgos y perjuicios, y la opinión de Moore prescribiría una adhesión rígida a reglas que nadie defendería.

Surge otra dificultad, estrechamente relacionada con la precedente, cuando Mill intenta aclarar el caso de una regla que debemos observar cuando provendría mayor bien de otra acción. 'Puede sostenerse que es expeditivo para un objeto inmediato, para algún propósito temporal, violar una regla cuya observación es conveniente en grado superior'. Así, 'a menudo podría ser conveniente decir una mentira para conseguir algo ventajoso para nosotros o para los demás' (p. 21). Pero .Mill alega que, de hecho, decir la mentira en tal caso no tendría mejores resultados que decir la verdad. Ha puesto a buen recaudo su argumentación llamando 'temporales' e 'inmediatos' a los buenos resultados de decir la mentira. Dice que, a la larga, decir la verdad tendrá más buenos resultados por dos razones: 'por cuanto que el cultivo de un sentimiento delicado respecto de la veracidad en nosotros es de los más proficuos, y el debilitamiento de tal sentido es muy nocivo; cosas éstas que pueden ser instrumento de nuestra conducta; y por cuanto cualquier desviación de la verdad, incluso preterintencional, inficiona en gran manera la probidad de la aserción humana en general' (p. 21). Ahora bien, el asunto principal que hay que hacer notar es que Mill para mientes en las consecuencias de decir esta verdad particular en este instante, y no en las consecuencias de decir la verdad en general.

Vale la pena advertir quizá que los dos argumentos en sí son inconcluyentes, puesto que son los que suelen emplear los utilitaristas del tipo ortodoxo o nourmsoniano para explicar por qué se ha de cumplir una regla en ciertas ocasiones cuando mejor bien redundaría quebrantándola. Cumplir la regla produ-

cirá bien a largo plazo por dos medios: (1) reforzando en el agente el hábito de cumplir la regla; (2) fomentando la seguridad que otros pondrán en guardarla. Discutiré estos argumentos en orden inverso por razones que aparecerán en su explanación.

Ross presentó una dificultad vital contra el argumento del 'fomento de la seguridad'. Si mi quebrantamiento de la regla no es conocido por nadie más, la seguridad general respecto de la ley queda intacta. En *The Right and the Good*, Ross ilustró este punto mediante lo que el señor Nowell-Smith ha llamado 'moralidad de la isla desierta' (*Ethica*, p. 240). Esto no está bien, pues Ross, en su último libro, *Foundations of Ethics*, trae un ejemplo simple de la vida real. Es importante notar que abundan los ejemplos de la vida real y son fáciles de hallar. En mi artículo de *Mind* (abril, 1939), titulado 'Punishment', que trata por entero de esta distinción entre utilitarismo ortodoxo y del tipo de Urmson, y de lo que ahora nos ocupamos también, cité dos casos vividos por mí y hablé de otro en 'Moral Rules' (*Proceedings of the British Academy*, 1953). Como se trata de un punto vital, brindo otro aquí. Un ex alumno mío era secretario de un hombre muy rico. Su patrón le había ordenado arrojar al cesto de los papeles, sin darles respuesta, todas las cartas petitorias. Era liberal con las causas que él había escogido, pero le era imposible averiguar la buena fe de cada petición. Este rico tenía también la costumbre de dejar montones de billetes en los bolsillos de sus trajes. Antes de mandar la ropa a la tintorería, su secretario revisaba los bolsillos y entregaba a su dueño los billetes encontrados. Este los volvía a poner en los bolsillos, sin contarlos. Una mañana en que no tenía qué hacer, el secretario se puso a leer las cartas petitorias por pura curiosidad, y entre ellas encontró una que tenía razón. Momentos antes había encontrado un montón de billetes en un bolsillo de una chaquetilla. Me contó que había dudado si sacar cinco billetes y enviárselos al firmante. 'El amo no lo habría descubierto

nunca.' Le pregunté si lo había hecho, pero me replicó: 'No, no era dinero mío.' No se trata de una razón utilitarista y, en particular, el hecho de que su patrón no lo hubiera sabido nunca invalida el argumento de 'fomento de la seguridad'. Mas cabe decir que había una persona que lo habría sabido, esto es, el secretario mismo, y aquí el utilitarista retornará al otro argumento. De haber enviado el dinero, el secretario habría debilitado su tendencia a no tomar lo que es ajeno y en otras ocasiones este debilitamiento habría tenido malos resultados. Pero tampoco vale este argumento. Para un utilitarista las reglas secundarias no se han de aplicar sin excepción y, por tanto, no se han de adquirir hábitos rígidos. El siguiente diálogo entre jugadores de bridge ilustrará la falacia. Soy el tercer jugador de la primera baza; el segundo jugador ha jugado ases; yo tengo rey. Recuerdo que se me ha dicho que el tercer jugador debe sacar alto. Susurro a mi mentor, que está detrás de mí: ¿Qué tiro? Me dice: 'Rey.' Replico: 'De nada va a servir, puesto que han jugado ases.' 'No importa; tú debes sacar el rey, pues de otra manera debilitarías tu tendencia a jugar alto al ser tercero.' '¿Pero se trata de una regla absoluta?' 'No, existen excepciones.' '¿Cuáles?' 'Cuando no lleva a nada tirar alto.' 'Pero ahora estamos en ese caso.' 'No importa; no debes debilitar tus buenos hábitos.'

Hay un paralelo interesante con este último punto en la manera como Mill trata los derechos. En su ensayo 'On Liberty' sostiene que no se debe impedir a nadie publicar sus opiniones científicas. Se apoya en que su opinión puede ser verdadera o parcialmente verdadera, en el cual caso será útil que sea conocida. Incluso si es falsa servirá para que quienes sustentan la opinión verdadera se alerten e impidan que la verdadera opinión quede en dogma muerto. El punto de especial interés aquí es el reconocimiento de que alguien puede decir que se tiene el derecho de publicar las opiniones científicas, incluso si de ello no resulta ningún beneficio. Su comentario establece:

'Es apropiado decir que declino cualquier ventaja que pudiera derivarse en favor de mi argumentación sobre la idea del derecho abstracto como algo independiente de la utilidad.' Se puede suponer que admite tal ventaja. Mas continúa: 'Considero la utilidad como apelación última en todas las cuestiones éticas, pero se trata de la utilidad en el más amplio sentido, fundada en los intereses permanentes del hombre como ser inteligente' (Everyman Edition, p. 74). Acude aquí, como en la argumentación referente a decir la verdad, a los resultados mediatos de la publicación del caso anterior. Ahora me ha caído en las manos una pequeña revista dedicada a defender que la tierra es plana. Es muy difícil mantener que ésta sea toda la verdad. Esa parte de verdad que puede decirse que ella contiene (que una minúscula parte de la superficie terrestre en casi plana) pertenece ya a la opinión ortodoxa. Y es difícil creer que la publicación de esta pequeña revista logre mantener en pie al astrónomo Royal. Con todo, la opinión de la mayoría se opondría a que se suprimiera esa publicación. Mas no es preciso denominar a esto derecho abstracto (o derecho autoevidente o natural). Podemos decir que *generalmente* es útil observar esta regla y aplicarla en todos los casos, aunque en algunas ocasiones no se derive algún bien de su aplicación. Esta sería la interpretación de Urmson, pero no parece que sea la argumentación de Mill.

Este artículo no ha tratado de los méritos rivales de los dos tipos de utilitarismo. Examiné ese aspecto en mis artículos sobre 'Punishment' ('El castigo') (1939) y sobre 'Moral Rules' ('Reglas morales') (1953), arriba citados. He tomado el texto de Mill con el fin de esclarecer las distinciones entre ambos.

Es interesante que en un artículo titulado 'Two Concepts of Rules' ('Dos conceptos de reglas') (*Philosophical Review*, vol. LXIV, enero de 1955), el señor J. B. Rawls trate el mismo tema e ilustre sus tesis remitiéndose a otro gran utilitarista, John Austin. Muestra de manera convincente que éste, en sus

Lectures on Jurisprudence (vol. I, p. 116), plantea muy claramente la interpretación que Urmson da del utilitarismo. Pero cuando pasa a discutirlo y defenderlo, se escurre hacia la interpretación ortodoxa, como he tratado de demostrar que hace Mill en su ensayo.

XI

DOS CONCEPTOS DE REGLAS [1]

John Rawls

De *Philosophical Review*, Vol. 64 (1955), pp. 3-32. Reimpreso con la venia del autor y de *Philosophical Review*.

En este artículo quiero esclarecer la importancia de la distinción entre justificar una práctica [2] y justificar una acción particular que cae dentro de ella, y deseo explicar la base lógica de esta distinción y por qué es posible preterir su trascendencia. Si bien se ha efectuado frecuentemente tal distinción [3] y aho-

1 Es una versión revisada de la disertación tenida en el Harvard Philosophy Club, el 30 de abril de 1954. [Lo discute H. J. McCloskey en 'An Examination of Restricted Utilitarism'. *Philosophical Review* (1957) y D. Lyons, *Forms and Limits of Utilitarism* (Clarendon Press, Oxford, 1965). El propio Rawls explica su posición en 'Justice as Fairness', *Philosophical Review* (1958), nota a la p. 168. E.]

2 Empleo la palabra 'práctica' en todo este artículo como una especie de tecnicismo que significa cualquier forma de actividad especificada por un sistema de reglas que define oficios, incumbencias, jugadas, castigos, defensas, etc. y que da su estructura a la actividad. Como ejemplos, piénsese en los juegos y en los rituales, en los juicios y en los parlamentos.

3 Esta distinción es fundamental en la discusión que Hume hace de lo que es la justicia en *A Treatise of Human Nature*, libro III, parte II, especialmente secs. 2-4. Se plantea claramente en la segunda conferencia de John Austin de *Lectures on Jurisprudence* (4.ª ed., Londres, 1873), i, 116 *ss.* (1.ª ed., 1832). Se puede alegar también que J. S. Mill la dio por sentada en *Utilitarism;* a este respecto, cf. J. O. Urmson, 'The Interpretation of the Moral Philosophy of J. S. Mill', *Philosophical Quarterly*, vol. iii (1953). Además

ra se está convirtiendo en tópico, queda todavía la tarea de explicar la tendencia, sea a soslayarla por entero, sea a menospreciar su importancia.

Para poner de manifiesto la importancia de tal distinción, defenderé el utilitarismo frente a aquellas objeciones que tradicionalmente se han dirigido contra él en conexión con el castigo y la obligación de cumplir las promesas. Espero demostrar que si alguien hace uso de tal distinción, será posible plantear el utilitarismo de modo que se expliquen mejor los juicios morales que consideremos, en comparación a lo que parecerían admitir las objeciones tradicionales [4]. Así, pues, mostraré la importancia de la distinción por el flanco que refuerza el punto de vista utilitarista, independientemente de si éste es del todo defendible o no.

Para explicar cómo puede preterirse la trascendencia de la distinción, discutiré dos conceptos de reglas. Uno de éstos entraña la importancia de distinguir entre la justificación de una regla o práctica y la justificación de una acción particular que caiga bajo ella. La otra concepción deja claro por qué se ha de hacer esa distinción y cuál es su base lógica.

de los argumentos dados allí por Urmson existen varios planteamientos claros de la distinción en *A System of Logic* (8 ed.; Londres, 1872). Libro VI, cap. xii, pars. 2, 3, 7. Esta distinción es importante en el artículo de J. D. Mabbott, 'Punishment', *Mind*, vol. xlviii (abril, 1939). Más recientemente, S. E. Toulmin, ha dado particular realce a la distinción en *The Place of Reason in Ethics* (Cambridge, 1950), ver esp. cap. xi, donde representa papel especial en su explicación del razonamiento moral. Toulmin no explica la base de la distinción ni cómo se pasa por alto su importancia, lo que trato de hacer aquí, y en la recensión de su libro (*Philosophical Review*, vol. lx [octubre 1951]), como muestran algunas de mis críticas no logré entender su fuerza. Ver también H. D. Aiken, 'The Levels of Mortal Discourse', *Ethics*, vol. lxii (1952), A. M. Quinton, 'Punishment', *Analysis*, vol. xiv (junio, 1954), y P. H. Nowell-Simth, *Ethics* (Londres, 1954), pp. 236-239, 271-273.

[4] Sobre el concepto de explicación, ver el artículo del autor *Philosophical Review*, vol. lx (abril 1951).

I

El sujeto del castigo, en el sentido de adjudicar penas legales a la violación de las reglas legales, ha constituido desde siempre cuestión moral batallona [5]. Ello no se debe a que la gente esté en desacuerdo sobre si es justificable o no la punición; la mayoría concede que, libre de ciertos abusos, es una institución aceptable. Pocos han sido los que la han rechazado por entero; lo que más bien sorprende, dadas las cosas que se pueden decir en su contra. La dificultad está en la justificación del castigo; a este efecto, los filósofos morales han esgrimido diversos argumentos, aunque hasta ahora ninguno do ellos ha ganado algún género de aceptación general; no hay justificación cabal para aquellos que detestan el castigo. Espero demostrar que el empleo de la distinción antes citada nos permitirá plantear el punto de vista utilitarista de modo que satisfaga los puntos válidos de sus críticos.

Para nuestros propósitos cabe decir que hay dos justificadores del castigo. Lo que podríamos llamar punto de vista retributivo establece que el castigo se justifica sobre la base de que las malas acciones merecen castigo. Está de acuerdo con la moral que alguien que hace el mal sufra en proporción con la maldad cometida. Que un criminal haya de ser castigado se sigue de su culpabilidad, y la severidad del castigo apropiado dependerá de la depravación de su acto. La situación cuando el malhechor sufre castigo es mejor moralmente que cuando no lo recibe, y es mejor independientemente de las consecuencias que se puedan seguir de castigarlo.

5 Mientras se corregía este artículo, apareció el de Quinton; nota 2 *supra* [nota 3, p. 205 de este volumen. E.]. Hay distintos aspectos que asemejan su artículo y el mío. Con todo, como considero algunas cuestiones ulteriores y me apoyo en argumentos algo distintos, he mantenido la discusión del castigo y de las promesas como dos casos-prueba del utilitarismo.

Lo que podemos denominar punto de vista utilitarista, fundado sobre el principio de que lo pasado es pasado y que las consecuencias futuras importan para las decisiones que se hayan de tomar, señala que el castigo se justifica sólo por referencia a las consecuencias probables de mantenerlo como uno de los instrumentos del orden social. Los errores cometidos en el pasado, como tales, no son consideraciones pertinentes que nos permitan decidir qué se ha de hacer. Si se puede demostrar que el castigo promueve efectivamente el interés de la sociedad, es justificable; de otra manera, no lo es.

He planteado de manera algo burda estos dos puntos de vista contrapuestos, para que se vea mejor la contención que existe entre ellos. Uno palpa la fuerza de *ambas* argumentaciones y se pregunta cómo es posible reconciliarlas. De mis observaciones introductorias se deduce que la resolución que voy a proponer consiste en que, en este caso, se ha de distinguir entre justificar una práctica como sistema de reglas que se pueden aplicar e imponer, y justificar una acción particular que cae bajo esas reglas. Los argumentos utilitaristas valen con cuestiones en torno a las prácticas, mientras que los argumentos retributivos se circunscriben a la aplicación de reglas particulares a casos particulares.

Aclararemos mejor esta distinción imaginando cómo un padre puede responder a su hijo. Supongamos que éste le pregunta: '¿Por qué ayer metieron en la cárcel a *J*?' El padre responde: 'Porque asaltó el banco de *B*. Se le juzgó debidamente y se le halló culpable; por eso lo pusieron ayer en la cárcel.' Pero supongamos que el hijo ha preguntado algo distinto, a saber: '¿Por qué unos ponen en la cárcel a otros?' Entonces el padre puede responder: 'Para proteger a los buenos de los malos' o 'Para impedir que haya gente que haga cosas que nos perjudicarían a todos, pues si no fuera así no podríamos ir a dormir

por la noche ni dormir en paz'. Hay aquí dos preguntas harto distintas. Una de ellas hace hincapié en el nombre propio: pregunta por qué se castigó a *J* y no a otro, o por qué se le castigó. La otra pregunta se refiere a por qué poseemos instituciones de castigo, por qué la gente castiga en vez, digamos, de perdonarse mutuamente.

Así, el padre dice que, en efecto, se castiga a un hombre determinado, y no a otro, porque es culpable, y lo es porque quebrantó la ley (tiempo pretérito). A su manera de ver, la ley mira hacia atrás, el juez mira hacia atrás y el jurado también mira hacia atrás, y se le impone una sanción por algo que cometió. Que se deba castigar a alguien y cuál es el castigo que se impondrá se estipula tras demostración de que quebrantó la ley y que ésta impone tal sanción por haber sido violada.

Por otra parte, tenemos la institución del castigo en sí, y recomendamos y aceptamos los distintos cambios que se le hagan porque el legislador (ideal) y aquellos a quienes se aplica la ley, cual parte de un sistema impuesto imparcialmente en cada caso que le corresponda, piensan que, a la larga, tendrá la consecuencia de fomentar los intereses de la sociedad.

Se puede decir, por tanto, que juez y legislador están en posiciones distintas y que miran en direcciones diferentes: uno hacia el pasado, el otro hacia el futuro. La justificación de lo que hace el juez, *en cuanto* juez, suena como punto de vista retributivo; la justificación de lo que el legislador (ideal) hace, *en cuanto* legislador, suena a punto de vista utilitarista. Así, las dos maneras tienen su razón (tal es como debe de ser, puesto que en un lado y otro de la argumentación ha habido personas inteligentes y sensatas). La confusión que se tiene al principio desaparece una vez se ve que esta manera de considerar las cosas se aplica a personas que efectúan distintos oficios con distintos deberes y están situadas dife-

rentemente con respecto al sistema de reglas que constituye la ley criminal [6].

Se podría decir, sin embargo, que la mira utilitarista es más fundamental, puesto que se aplica a oficio más fundamental, ya que el juez ejecuta la voluntad del legislador hasta donde puede determinarla. Una vez que el legislador decide tener leyes y aplicar sanciones por su violación (según sean las cosas, tienen que existir tanto la ley como el castigo) se erige una institución que contiene una concepción retributiva de los casos particulares. Es parte del concepto de ley criminal, como sistema de reglas, que la aplicación e imposición de éstas en casos particulares se han de poder justificar por argumentos de carácter retributivo. La decisión de emplear la ley y no otro mecanismo de control social, y la resolución acerca de cuáles han de ser esas leyes y qué sanciones se han de asignar, puede establecerse en argumentaciones utilitaristas, pero si se decide tener leyes, entonces se ha resuelto sobre algo cuyo funcionamiento en los casos particulares es retributivo por su forma [7].

La respuesta, pues, a la confusión engendrada por los dos puntos de vista del castigo es muy simple: se distinguen dos oficios, el del juez y el del legislador, y se distinguen sus distintas situaciones con respecto al sistema de reglas que constituyen la ley; entonces se advierte que los diferentes tipos de consideraciones, que de ordinario se presentarían como razones de lo que se lleva a cabo bajo la cubierta de estas funciones, se pueden emparejar con las justificaciones conflictivas del castigo. Se reconcilian los dos puntos de vista por el procedimiento de aplicarlos a diferentes situaciones sancionado por el tiempo.

6 Adviértase el hecho de que para los distintos oficios cuadran distintas clases de argumentaciones. Una manera de señalar las diferencias entre las teorías éticas es considerarlas como explicaciones de las razones que fundan los diferentes oficios.

7 A este respecto, ver Mabbott, *op. cit.*, pp. 163-164.

¿Pero es tan simple esto? Bien, en la respuesta se ha de tener presente el propósito aparente de cada lado. Quien defienda el punto de vista retributivo ¿ha de abogar necesariamente por la maquinaria legal, como *institución* cuyo propósito esencial es instaurar y preservar la correspondencia entre la torpeza moral y el sufrimiento? No, sin duda [8]. En lo que los retribucionistas han insistido con razón es en que nadie puede ser castigado, a menos que sea culpable, o sea, a menos que haya quebrantado la ley. Su crítica fundamental de la razón utilitarista es que, según la interpretan, sanciona que se castigue a una persona inocente (si se le puede llamar castigo) en aras de la sociedad.

Por otra parte, aceptan los utilitaristas que el castigo se ha de imponer sólo por la violación de la ley; consideran que esto se sobreentiende por el mismo concepto de castigo. La tesis utilitarista se refiere a la institución como sistema de reglas; el utilitarismo intenta limitar su empleo declarándola justificable sólo si se puede demostrar que secunda de manera efectiva el bien de la sociedad. Históricamente, es una protesta contra el uso indiscriminado e inefectivo de la ley criminal [10]. Trata de disuadirnos de asignar a las

[8] A este respecto ver Sir David Ross, *The Right and the Good* (Oxford, 1930), pp. 57-60.

[9] Ver la definición de castigo que Hobbes trae en *Leviathan*, cap. xxviii, y la definición de Bentham en *The Principle of Morals and Legislation*, cap. xii, par. 36, cap. xv, par. 28, y en *The Rationale of Punishment* (Londres, 1830), libro 1, cap. i. Podrían concordar con Bradley en que: 'El castigo es castigo sólo cuando se merece. Se paga la penalidad porque se debe y por ninguna otra razón, y si se inflige castigo por alguna otra razón, sea cual sea, y no porque esté merecido por haber hecho el mal, es una burda inmoralidad, una injusticia clamante, un crimen abominable y no lo que pretende ser'. *Ethical Studies* (2.ª ed., Oxford, 1927), páginas 26-27. Ciertamente, por definición no es lo que pretende ser. El inocente sólo puede ser castigado por error; el 'castigo' deliberado del inocente comporta fraude necesariamente.

[10] Cf. Leon Radzinowicz, *A History of English Criminal Law: The Movement for Reform 1750-1833* (Londres, 1948, esp. cap. xi sobre Bentham).

instituciones penales la tarea impropia, si no sacrílega, de equiparar el sufrimiento con la torpeza moral. Al igual que otros, los utilitaristas quieren que las instituciones penales estén de tal manera dispuestas que, hasta donde sea humanamente posible, sólo quienes quebranten la ley tengan que habérselas con ella. Defienden que ningún oficial debería tener poder a discreción para infligir castigos cuando lo considerara beneficioso para la sociedad, pues, según los utilitaristas, una institución que dé pie a tales cosas no tiene justificación [11].

La manera aquí sugerida para reconciliar las justificaciones retributiva y utilitarista del castigo parece dar razón de lo que ambos bandos han querido decir. Hay, sin embargo, otras dos cuestiones, a las que dedicaré lo que resta de esta sección.

Primero, ¿no será inconveniente, para que los retribucionistas acepten la reconciliación, la diferencia de opinión respecto del criterio apropiado de lo que es ley justa? ¿No pondrán en duda que si se aplica como criterio el principio utilitarista, se siga que quienes hayan quebrantado la ley sean culpables, de modo que se satisfaga su alegato de que quienes se castiguen lo merezcan? Para responder a esta dificultad, supongamos que las reglas de la ley criminal se justifican según las bases utilitaristas (sólo se puede hacer responsable al utilitarista de leyes

11 Bentham trata de cómo, en correspondencia a la provisión punitiva de la ley criminal, hay otra provisión que está en antagonismo con ella y que merece nombre lo mismo que la punitiva; la denomina, como se podía esperar, *anetiosóstica*, y dice de ella: 'El castigo de la culpa es el objeto de la primera; la preservación de la inocencia, el de la segunda'. En la misma conexión afirma que nunca es conveniente dar al juez la opción de decidir si un ladrón (esto es, una persona a la que cree ladrón, puesto que la creencia del juez es en torno a lo que siempre ha de girar la cuestión) ha de ser ahorcado o no, por lo que la ley prescribe la provisión: 'El juez no hará que se ahorque a un ladrón a menos que sea debidamente convicto y sentenciado en el curso de la ley' (*The Limits of Jurisprudence Defined*, ed. C. W. Everett [Nueva York, 1945], pp. 238-239).

que se ajusten a su criterio). Se sigue entonces que las acciones que la ley criminal especifica como ofensas son de tal manera que, si se toleraran, esparcirían terror y alarma por la sociedad. Consiguientemente, los retribucionistas sólo pueden denegar que quienes son castigados merecen serlo, si niegan que tales acciones son malas. Pero no lo negarán.

La segunda cuestión es sobre si el utilitarismo justifica demasiado. Se nos imagina como una máquina de justificación que, si se amañara convenientemente, podría emplearse para justificar instituciones crueles y arbitrarias. Los retribucionistas conceden, por sentado, que los utilitaristas *pretenden* reformar la ley y hacerla más humana; que no *quieren* justificar cosas como la sanción del inocente y que pueden apelar al hecho de que el castigo presupone culpabilidad, si se entiende por castigo una institución que grava con penalidades la infracción de las reglas legales, y que, por tanto, es lógicamente absurdo suponer que los utilitaristas, al justificar el *castigo*, justifiquen también el castigo (si así lo podemos llamar) del inocente. La verdadera cuestión, empero, es si el utilitarista, al justificar el castigo, no ha empleado argumentos que lo comprometen a aceptar la imposición de sufrimientos a personas inocentes, si es para el bien de la sociedad (llámesele o no castigo). De una manera más general, ¿no está obligado el utilitarista, en principio, a aceptar muchas prácticas que como persona moralmente sensata no ha de querer aceptar? Los retribucionistas se inclinan a sostener que no es posible impedir que el principio utilitarista justifique demasiado, a menos que se le añada un principio que distribuya ciertos derechos entre los individuos. Entonces el criterio enmendado no es el mayor beneficio de la sociedad *simpliciter*, sino el mayor beneficio de la sociedad, con la reserva de que no se han de violar derechos de nadie. Ahora bien, si soy de la opinión de que los utilitaristas clásicos propusieron un criterio de este género más complicado, no es mi intención

dilucidar aquí este asunto [12]. Lo que quiero mostrar es que existe *otro* medio para impedir que el principio utilitarista justifique demasiadas cosas o, siquiera, para conseguir que sea menos probable que las justifique; a saber, planteando el utilitarismo de guisa que comprenda la distinción entre la justificación de una institución y la justificación de una acción particular que caiga dentro de ella.

Empero definiendo así la institución del castigo: se dice que una persona sufre castigo cuando legalmente se le priva de algunos de los derechos normales de todo ciudadano, sobre razón de que ha violado alguna regla de la ley, tras haberse probado la violación por juicio, a tenor del debido proceso legal, habida cuenta de que la privación se efectúe por las autoridades legalmente reconocidas del estado, que la regla de la ley especifique claramente tanto la ofensa como la penalidad consiguiente, que los tribunales estipulen estrictamente los estatutos y que el estatuto esté registrado con anterioridad al tiempo de la ofensa [13]. Esta definición específica qué es lo que entenderé por punición. La cuestión es si las argumentaciones utilitaristas justifican instituciones que difieren notablemente de ésta, y que pueden considerarse crueles o arbitrarias.

Se responderá mejor a esta cuestión, según creo, considerando una acusación particular. Veamos lo siguiente de Carritt:

> ...el utilitarista debe sostener que será justo que inflijamos daño siempre y sólo para impedir daño peor o atraer mayor felicidad. Esto, pues, es todo lo que necesitamos considerar en el llamado castigo, que ha de ser puramente preventivo. Pero si se generaliza algún tipo de crimen cruel y es imposible aprehender a ninguno de los facinerosos, puede ser altamente expeditivo, por ejemplo, ahorcar a un inocente si se pudiera maquinar contra él algún cargo, de modo que a la vista de todos pasara por culpable; en

12 Por utilitaristas clásicos entiendo a Hobbes, Hume, Bentham, J. S. Mill y Sidgwick.

13 Hobbes menciona todas estas características del castigo; cf. *Leviathan*, cap. xxviii.

realidad esto no sería dechado de 'castigo' utilitarista, exclusivamente porque la víctima no habría sido felón que fuera a cometer tal fechoría en el futuro; en los demás aspectos sería perfectamente disuasor y, por tanto, para bien.

Carritt trata de demostrar que existen ocasiones en que la argumentación utilitarista justificaría emprender una acción que se condenaría en general y que, por tanto, el utilitarismo se excede en justificar. Pero la falla del argumento de Carritt yace en el hecho de que no hace distinción entre la justificación del sistema general de reglas, que constituye las instituciones penales, y la justificación de aplicaciones particulares de esas reglas a casos particulares, por parte de los distintos oficiales a quienes compete administrarlas. Esto se hace del todo claro cuando se pregunta quiénes son el 'nosotros' de que habla Carritt. ¿Quién es aquel que dispone de una clase de autoridad absoluta en ocasiones particulares para decidir que se 'castigue' a un inocente, si se puede convencer a los demás de que es culpable? Dicha persona ¿es el legislador, el juez o el cuerpo de los ciudadanos privados, o quién? Es del todo imprescindible saber quién decide en tales cuestiones y con qué autoridad, pues todo esto ha de constar en las reglas de la institución. Si no se saben estas cosas, no se conocerá cuál es la institución cuya justificación se pone en duda, y como el principio utilitarista se refiere a la institución, no se sabe tampoco si está justificada según la mira utilitarista, o no lo está.

Una vez entendido esto, queda claro cuál ha de ser el despliegue frente al argumento de Carritt. Se ha de describir más detalladamente cuál es la *institución* que sugiere su ejemplo, y entonces preguntarse si es probable que poseer tal institución sea provechoso, a la larga, para la sociedad o no. No se ha de contentar uno con el pensamiento vago de que, cuando se trata de *este* caso, sería buena cosa si *alguien*

14 *Ethical and Political Thinking* (Oxford, 1947), p. 65.

hiciera algo, aunque tuviera que pagar algún inocente.

Trátese de imaginar, pues, una institución (que podríamos denominar 'telismo' *) que fuera tal que los funcionarios nominados tuvieran autoridad de disponer un juicio para la condenación de un inocente, siempre que lo consideraran oportuno porque redundara en pro de los intereses de la sociedad. La discreción de tales funcionarios, sin embargo, estaría limitada por una regla que estatuiría que no podrían condenar a un inocente a sufrir tal prueba a menos que, a la sazón, hubiera una ola de desmanes similares a aquéllos de que le acusan y por los que le 'telizan'. Podemos imaginar que los funcionarios que tienen la autoridad discrecional son los jueces de los tribunales más altos, en consulta con el jefe de la policía, con el ministro de la justicia y con un comité de la legislatura.

Cuando uno se percata de que ha de instaurar una *institución*, se ve que los riesgos son muy grandes. Por ejemplo, ¿qué control tienen los funcionarios? ¿Cómo se determinará si sus acciones están autorizadas o no? ¿Cómo se han de limitar los riesgos provenientes de permitir tal impostura sistemática? ¿Cómo se ha de evitar él conceder a las autoridades algo que carezca de discreción, por lo que 'telicen' a quien quieran? Además de estas consideraciones, es obvio que los ciudadanos tendrán una actitud muy diferente hacia su sistema penal cuando se le yuxtaponga el 'telismo'. No sabrán a ciencia cierta si un individuo convicto ha sido castigado o 'telizado'. Se preguntarán si lo han de sentir o no; se preguntarán si alguna vez no les tocará el mismo sino. Si uno se imagina cómo funcionaría en realidad tal institución y los enormes riesgos que comportaría, parece claro que no sería de ningún provecho. No es posible que

* El autor ha inventado la palabra 'telishment' y el verbo 'to telish' (T.).

exista justificación utilitarista respecto de esta institución.

Sucede en general que si se dejan de lado las características definitorias del castigo, no queda más que una institución cuya justificación utilitarista es altamente dudosa. Una de las razones está en que el castigo funciona como una especie de sistema de precios: si se alteran los precios que se tienen para pagar por la ejecución de las acciones, surge un motivo para evitar unas acciones y hacer otras. Las características definitorias son esenciales si el castigo ha de operar así; por lo que una institución que carezca de esas características, *v. g.*, una institución que esté dispuesta de manera que 'castigue' al inocente, es como si tuviera un sistema de precios (si así vale llamarlo) en que los precios variaran al azar día a día y sólo se supieran luego de haber aceptado comprar el artículo [15].

Si se tiene cuidado de aplicar el principio utilitarista que autoriza acciones particulares, entonces hay

15 La analogía con el sistema de precios sugiere una respuesta a la cuestión sobre cómo las consideraciones utilitaristas garantizan que el castigo sea proporcional a la ofensa. Es interesante advertir que Sir David Ross, tras hacer la distinción entre justificar una ley penal y justificar su aplicación particular, y después de plantear que las consideraciones utilitaristas tienen amplio lugar para determinar lo primero, se abstiene de aceptar la justificación utilitarista del castigo, sobre las bases de que la justicia requiere que el castigo sea proporcional a la ofensa y que el utilitarismo es incapaz de dar razón de esto. Cf. *The Right and the Good*, pp. 61-62. No digo que el utilitarismo contenga este requisito, como podría desear Sir David, pero sucede, no obstante, que si se siguen las consideraciones utilitaristas, las penas serán proporcionales a las ofensas en este sentido: el orden de las ofensas, de acuerdo con su seriedad, puede equipararse con el orden de las penas de acuerdo con la severidad. También el nivel absoluto de las penas será tan bajo como sea posible. Esto se sigue de la presuposición de que la gente es racional (esto es, de que es capaz de tomar en cuenta los 'precios' que impone el estado sobre las acciones), de la regla utilitarista de que un sistema penal ha de dar motivo para preferir la ofensa menos seria y el principio de que el castigo como tal es un mal. Todo esto fue elaborado cuidadosamente por Bentham en *The Principles of Morals and Legislation*, caps. xiii-xv.

menos peligro de que se exceda justificando. El ejemplo de Carrit es admisible porque es indefinido y se concentra en el caso particular. Su argumentación se sostendrá en pie sólo si se puede demostrar que existen argumentos utilitaristas que justifican una institución cuyos oficios y poderes públicamente discrecionales son de tal cariz que permiten a los funcionarios ejercitar ese tipo de discreción en los casos particulares. Pero el requisito de tener que incorporar características arbitrarias en la práctica institucional desmejora mucho su justificación.

II

Consideraré ahora la cuestión de las promesas. La objeción que se hace al utilitarismo por lo referente a las promesas parece ser ésta: se cree que, desde el punto de vista utilitarista, cuando alguien hace una promesa, el único fundamento por el que ha de cumplirla, si la ha de cumplir, es que ajustándose a ella cooperará al mejor bien de todos. Así, cuando alguien pregunta: '¿Por qué he de cumplir *mi* promesa?', se entiende que la respuesta utilitarista será que, al actuar así en *este* caso, se obtendrán las mejores consecuencias. Y se dice con razón que esta respuesta choca con la manera como se considera la obligación de cumplir las promesas.

Es claro que a los críticos del utilitarismo no se les escapa que una de las defensas que se atribuyen a los utilitaristas se refiere a la práctica del cumplimiento de lo prometido [16]. En este sentido, se supone

16 Ross, *The Right and the Good*, pp. 37-39, y *Foundations of Ethics* (Oxford, 1939), pp. 92-94. No conozco a ningún utilitarista que haya empleado este argumento excepto W. A. Pickard-Cambridge en 'Two Problems about Duty', *Mind*, xli (abril 1932), 153-157, aunque el argumento va con la versión mooreana del utilitarismo en

que argumentan así: se ha de admitir que nuestro pensar respecto de las promesas es estricto, más estricto que cuanto pudiera inferirse de nuestro modo de ver. Pero cuando consideramos atentamente este asunto, es preciso siempre tomar en cuenta el efecto que nuestra acción tendrá en la práctica del cumplimiento de las promesas. Quien promete ha de sopesar no sólo los efectos de quebrantar su promesa, en el caso particular, sino también el efecto que tendrá sobre la propia práctica si se la quebranta. Puesto que la práctica es de gran valor utilitarista, y puesto que romper las promesas siempre la daña de manera seria, raramente se justificará que alguien quebrante sus promesas. Si se consideran nuestras promesas individuales en el contexto más vasto de la práctica de la promisión en sí, comprenderemos lo estricto de la obligación de cumplirlas. Existe siempre una consideración utilitarista muy fuerte en favor de cumplirlas, y ésta reforzará el consenso afirmativo cuando se pregunte si se han de cumplir o no, incluso cuando los hechos de un caso particular tomado en sí parezcan justificar su quebrantamiento. De esta guisa, damos razón del rigor con que vemos la obligación de cumplir las promesas.

Ross ha criticado esta defensa como sigue [17]: por grande que sea el valor de la práctica de la promisión según bases utilitaristas, tiene que haber algún valor que sea mayor y pensarse que es posible obtenerlo por violación de las promesas. Por tanto, puede existir un caso en el cual quien prometa alegue que romper la promesa hecha estaba justificado, porque conducía a una situación mejor en su totalidad, arguyendo así aparte de cuán nimia fuera la ventaja que recabara quebrantando la promesa. Si se quisiera disceptar con el prometiente, se defendería diciendo

Principia Ethica (Cambridge, 1903). Por lo que sé, no aparece en los utilitaristas clásicos, y si se interpreta correctamente su punto de vista, ello no se debe a casualidad.

17 Ross, *The Right and the Good*, pp. 38-39.

que su proceder fue mejor en su totalidad, vistas todas las consideraciones utilitaristas, que en este caso *comprenden* la importancia de la práctica. Ross considera que tal defensa es inaceptable. Creo que tiene razón, en cuanto que protesta contra la apelación a las consecuencias en general y sin mayor explicación. Con todo, es extremadamente difícil sopesar la fuerza del argumento de Ross. El caso descrito tiene cariz irreal y parece pedir más descripción. Uno se siente inclinado a pensar, o que tal caso pertenecería a una excepción definida por la práctica misma, contingencia en que no valdría apelar a las consecuencias en general en ese caso particular, o que las circunstancias eran tan peculiares que no tendrían lugar las condiciones que presupone la práctica. Pero Ross tiene razón en pensar que nos sorprende como algo errado que una persona defienda la ruptura de una promesa apelando de manera general a las consecuencias. El prometiente, desde luego, no tiene defensa utilitarista general: no es una de las defensas permitidas por la práctica de la promisión.

Ross trae dos argumentaciones más en contra [18]: en primer lugar, dice que se sobreestima el perjuicio que se causa a la práctica del prometer por una falla en cumplir una promesa. Quien no cumple una promesa mancilla su nombre, sin duda alguna; pero no está del todo claro que una promesa rota dañe siempre la práctica misma como para que dé razón del rigor en la obligación. En segundo lugar, y lo que creo más importante, se pregunta qué se ha de decir de una promesa que nadie sabe que ha sido pronunciada, excepto quien promete y quien recepta, como en el caso de la promesa que hace un hijo a su padre mo-

18 Ross, *ibid.*, p. 39. El caso de la promesa no pública vuelve a tratarse en *Foundations of Ethics*, pp. 95-96, 104-105. Ocurre también en Mabbott, 'Punishment', *op. cit.*, pp. 155-157, y en A. I. Melden, 'Two Comments on Utilitarism', *Philosophical Review*, lx (octubre 1951), 519-523, quien discute el ejemplo de Carritt en *Ethical and Political Thinking*, p. 64.

ribundo sobre el manejo de la hacienda [19]. En este tipo de caso, la consideración respecto de la práctica no tiene peso absolutamente sobre el prometiente, pero con todo se siente que esta forma de promesa obliga tanto como las demás. La cuestión del efecto que sobre la práctica tiene la ruptura de las promesas parece del todo irrelevante; la única consecuencia parece ser que se puede quebrantar la promesa sin riesgo de ser censurado, pero la obligación no parece disminuida en lo más mínimo. Puesto que es dudoso si el efecto sobre la práctica pesa siempre en un caso particular, no puede dar razón cierta sobre el rigor de la obligación cuando ese efecto no tiene lugar. Parece seguirse que la razón utilitarista de la obligación de cumplir las promesas no tiene prospectos de éxito.

Por lo que he dicho en conexión con el castigo, se puede prever lo que voy a decir acerca de estos argumentos y objeciones. No distinguen entre justificación de una práctica y justificación de una acción particular que pertenece a aquélla, con lo que caen en el error de suponer que el prometiente, al igual que el oficial de Carritt, tiene licencia para llevar a la práctica, sin restricción, consideraciones de tipo utilitarista para decidir sobre el cumplimiento de *su* promesa. Pero si se atiende a lo que es la práctica de la promisión, se verá —creo— que es de tal suerte que no permite al prometiente este tipo de discreción general. En efecto, el quid de la práctica es abdicar el título propio para actuar de acuerdo con las consideraciones utilitaristas y prudenciales, con el fin de consolidar el futuro y predisponer planes con antela-

19 El ejemplo de Ross se refiere simplemente a dos hombres que mueren solos y uno hace una promesa al otro. El ejemplo de Carritt (cf. n. 17 *supra)* [nota 1. E.] es de dos hombres que están en el Polo Norte. El ejemplo del texto es más realista y se asemeja al de Mabbott. Otro ejemplo es cuando alguien comunica algo confidencialmente a otro y luego muere. Tales casos no precisan ser 'argumentos de isla desierta' como Nowell-Smith parece creer (cf. su *Ethica,* pp. 239-244).

ción. Existen ventajas utilitaristas obvias en disponer de una práctica que deniegue al prometiente, como defensa, cualquier apelación general al principio utilitarista de acuerdo con el cual se pueda justificar la práctica. Nada hay ni contradictorio ni sorprendente en esto: se pueden dar razones válidas utilitaristas (o estéticas) en favor de que el ajedrez o el baseball están bien como están, o en pro de que se deberían cambiar en determinadas cosas, pero el jugador no puede apelar a tales consideraciones, en el juego, como motivos para proceder a su modo. Es error pensar que si se justifica la práctica según motivos utilitaristas, entonces quien promete ha de disponer de libertad total para emplear argumentos utilitaristas en decidir si ha de cumplir o no una promesa. La práctica prohíbe esta defensa general, y buen motivo tiene para hacerlo. Por tanto, lo que presuponen los anteriores argumentos —la idea de que en la mira utilitarista el prometiente está obligado si, y sólo sí, la aplicación del principio utilitarista a su propio caso muestra que cumplir la promesa es lo mejor en conjunto— es falso. El prometiente está obligado porque prometió; no depende de él juzgar el caso según lo merezca [20].

¿Quiere esto decir que en casos particulares no se puede deliberar si se ha de cumplir una promesa o no? Por supuesto que no. Pero proceder así equivale a deliberar si las distintas excusas, excepciones y defensas que se entienden por la práctica y constituyen parte importante de ella se aplican al propio caso [21]. Hay varias excusas para no cumplir las promesas, pero no hay ninguna según la cual, fundándose en motivos utilitaristas generales, el prometiente pueda pensar (verdaderamente) que en su totalidad su proceder es el mejor, aunque pueda tener la disculpa

20 Para una discusión de esto, ver H. Sidgwick, *The Methods* la importante discusión de Hume en *Treatise of Human Nature*, libro III, parte II, sec. 5, y también sec. 6, par. 8.

21 Para una discusión de esto, ver H. Sidgwick, *The Methods of Ethics* (6.ª ed., Londres, 1901), libro III, cap. vi.

de que las consecuencias de cumplir la promesa habrían sido en *extremo* serias. Si bien hay aquí sobradas complejidades para poder considerar todos los detalles necesarios, se puede ver que no cabe excusa general si se pregunta lo siguiente: ¿qué se diría de alguien que, al preguntársele por qué no se atuvo a la promesa, replicara simplemente que lo mejor en general fue quebrantarla? Suponiendo que su respuesta fuera sincera y que su creencia fuera razonable (es decir, sin considerar que estuviera equivocado), creo que uno se preguntaría si sabe qué significa 'prometo' (en las debidas circunstancias). Se diría de alguien que empleara esta excusa sin mayor explicación que no entiende qué defensas le permite la práctica que define lo que es una promesa. Si un niño echara mano de esta excusa, se le corregiría, pues es parte de cómo se nos inculca el concepto de promesa el corregir el empleo de tal excusa. La práctica caería por el suelo si aquella excusa fuera permitida.

No hay duda de que es parte del punto de vista utilitarista que toda práctica ha de admitir el descargo de que las consecuencias de atenerse a esa práctica habrían sido en extremo serias. Además, los utilitaristas se inclinarían a conceder alguna confianza en el buen sentido de la gente y que es preciso hacer concesiones en casos difíciles. Mantendrían que una práctica se justifica si sirve a los intereses de quienes la comparten, pues, como con cualquier conjunto de reglas, se sobreentiende que existe un trasfondo de circunstancias bajo las cuales es natural que se aplique, circunstancias que no es preciso —ni es posible— detallar. Si estas circunstancias cambian, entonces, aunque no haya regla alguna que dé razón del caso, a lo mejor todavía está de acuerdo con la práctica que alguien quede libre de la obligación. Pero este tipo de excusa permitido por la práctica no se ha de confundir con la opción general de sopesar cada caso particular sobre base utilitarista, que

los críticos del utilitarismo han considerado indispensable.

El óbice que se pone a la justificación que el utilitarismo permite respecto del castigo es que puede justificar demasiado. La cuestión referente a las promesas es diferente, pues se trata de cómo el utilitarismo justifica la obligación de cumplir las promesas. Uno siente que la obligación reconocida de cumplir con las promesas y el utilitarismo son incompatibles. Y sin duda lo son si se interpreta que el punto de vista utilitarista sostiene que cada individuo es completamente libre de medir cada acción particular según motivos utilitaristas generales. Pero ¿se ha de interpretar así el utilitarismo? Espero mostrar que, en los casos que he tratado, no se puede interpretar así.

III

Hasta aquí he tratado de mostrar la importancia de la distinción entre la justificación de una práctica y la justificación de una acción particular que cae bajo ella, indicando cómo se puede utilizar esta distinción para defender el utilitarismo contra dos objeciones tenaces. Puede sentirse la tentación de cerrar la discusión en este punto, diciendo que las consideraciones utilitaristas se han de entender como aplicables a prácticas del primer caso y no a las acciones particulares que caen bajo ellas, excepto hasta donde esas prácticas lo permiten. Podría alguien decir que, por esta forma modificada, se da mejor razón de las opiniones morales que hemos considerado, y dejar así la cosa. Pero detenerse aquí sería preterir la interesante cuestión sobre cómo es posible que se deje de apreciar la importancia de esta distinción, que más bien es obvia, y pueda darse por sentado que el utili-

tarismo tiene la consecuencia de que los casos particulares pueden decidirse siempre según principios utilitaristas generales[22]. Me parece que este error se debe a una concepción equivocada del *status* lógico de las reglas de las prácticas. Para demostrar este particular examinaré dos conceptos de reglas, dos modos de inserirlas dentro de la teoría utilitarista.

La concepción que entraña la trascendencia de la distinción recibirá aquí el nombre de mira sumaria. Considera así las reglas: se supone que cada persona decide qué ha de hacer en los casos particulares, aplicando el principio utilitarista; se supone, además, que las diferentes personas decidirán un mismo caso particular de la misma manera y que habrá recurrencias de casos similares a los que se decidieron previamente. Sucederá que, en casos de cierto tipo, se tomará la misma decisión, sea por la misma persona en diferentes ocasiones, o por distintas personas al mismo tiempo. Si ocurre un caso con la suficiente frecuencia, se supone que se formulará una regla que rija ese tipo de caso. He llamado a esta concepción mira sumaria porque las reglas se imaginan como sumarios de las decisiones pasadas, a las que se llegó

22 Hasta donde me es dado conocer, no es sino con Moore cuando esta doctrina se plantea expresamente de esta manera. Ver, por ejemplo, *Principia Ethica*, p. 147, donde se dice que la proposición 'Estoy obligado moralmente a realizar esta acción' es idéntica que la proposición '*Esta* acción producirá la mayor cantidad posible de bien en el Universo' (cursivas mías). Es importante recordar que aquellos a quienes denomino utilitaristas clásicos estaban muy interesados por las instituciones sociales. Estaban entre los economistas guías y entre los teóricos políticos de sus días y no era raro que fueran reformadores preocupados por los asuntos prácticos. Históricamente, el utilitarismo va de consuno con una visión coherente de la sociedad y no es simplemente una teoría ética y, mucho menos, un conato de análisis filosófico en el sentido moderno. El principio utilitarista se consideró y utilizó como criterio para juzgar las instituciones sociales (prácticas), y como base para urgir las reformas. No está claro, por tanto, hasta qué grado se ha de enmendar el utilitarismo de forma clásica. Para una discusión sobre el utilitarismo como parte integral de una teoría de la sociedad, ver L. Robbins, *The Theory of Economic Policy in English Classical Political Economy* (Londres, 1952).

por aplicación *directa* del principio utilitarista a los casos particulares. Las reglas se consideran como informes de que cierto tipo de casos se han resuelto apropiadamente de determinada manera, sobre *otras* bases (aunque, es claro, no lo *dicen*).

Hay varias cosas que advertir respecto de esta guisa de inserir reglas en la teoría utilitarista [23].

23 Esta nota se ha de leer después de la sec. 3 y presupone lo que allí he dicho. Se trata de unas cuantas referencias a asertos de utilitaristas importantes de la mira sumaria. En general, parece que cuando trataban las características lógicas de las reglas, fue la mira sumaria la que prevalecía y era lo típico de cómo hablaban acerca de las reglas morales. Cito un conjunto algo largo de pasajes de Austin, como ilustración cabal.

John Austin en sus *Lectures on Jurisprudence* contradice la objeción de que decidir de acuerdo con el principio utilitarista caso por caso sea impráctico, afirmando que es una interpretación equivocada del utilitarismo. Según el punto de vista utilitarista, '...nuestra conducta se ha de conformar a las *reglas* inferidas de las tendencias de las acciones, pero no se ha de determinar acudiendo directamente al principio de la utilidad general. La utilidad ha de ser la piedra de toque de nuestras acciones en última instancia, no de manera inmediata; ha de ser la piedra de toque inmediata de las reglas a las que se ha de conformar nuestra conducta, pero no la piedra de toque inmediata de las acciones específicas o individuales. Nuestras acciones se han de cortar según la utilidad; nuestra conducta, según nuestras reglas' (vol. 1, p. 116). Respecto de cómo se decide sobre la tendencia de una acción, dice: 'Si queremos probar cuál es la tendencia de un acto individual o específico, no debemos contemplar el acto como si fuera solo o estuviera aislado, sino que hemos de ver la clase de actos a que pertenece. Debemos suponer que los actos de esa clase son hechos u omitidos generalmente, y considerar su efecto probable sobre la felicidad o bien generales. Tenemos que adivinar las consecuencias que se seguirían si esa clase de actos fuera general, así como las consecuencias que se seguirían si se omitieran de ordinario. Entonces hemos de comparar tales consecuencias en lo positivo y negativo y ponderar sobre qué lado pesa el *platillo* de la ventaja... Si comprobamos verdaderamente la tendencia de un acto específico o individual, comprobamos la tendencia de la clase a que pertenece el acto. La conclusión *particular* que extraemos, respecto de ese acto individual, implica una conclusión *general* que abarca todos los actos similares .. A las reglas así colegidas y almacenadas en la memoria se amoldará *inmediatamente* nuestra conducta, si ellas se ajustan verdaderamente a la utilidad' *(ibid.,* p. 117). Se puede pensar que Austin contesta a la objeción siguiendo la idea de la práctica de las reglas, y quizá fue esto lo que intentó. Pero no es claro que así lo haya hecho. La generalidad a que se refiere, ¿es

1. La razón de poseer reglas está en el hecho de que hay ciertos casos que tienden a recurrir y en que se resuelven los casos con tanta mayor facilidad si se dispone de resoluciones pasadas en forma de reglas. Si tales casos similares no volvieran a recu-

de tipo estadístico? Tal se infiere por la noción de tendencia; ¿o se refiere a la utilidad de establecer una práctica? No lo sé, pero sus observaciones subsiguientes parecen seguir la mira sumaria. Dice: 'Considerar las consecuencias específicas de los actos particulares o individuales, *raramente* [cursivas mías] sería consecuente con el principio último' (*ibid.*, p. 117). Pero ¿se ha de proceder así alguna vez? Continúa: '...admitido esto, la necesidad de detenerse a calcular, que supone la objección de la cuestión, es imaginaria. Prolongar cada acto o demorarlo con una conjetura y comparación de las consecuencias sería claramente *superfluo* [cursivas mías] y malintencionado. Sería claramente superfluo, por cuanto que el *resultado de ese proceso* [cursivas mías] quedaría incorporado en una regla conocida. Sería claramente malintencionado, por cuanto el *verdadero* resultado se expresaría por esa regla, mientras que el proceso probablemente quedaría defectuoso si se efectuara según el acicate de la ocasión' (*ibid.*, pp. 117-118). Continúa: 'Si no se *generalizaran* nuestra experiencia y observación de los particulares, de poco aval nos serían nuestra experiencia y observación de los particulares en la *práctica*... Las inferencias que acuden a nuestras mentes, por la experiencia y observación repetidas, se concluyen en *principios* o se comprimen en *máximas*, que llevamos encima listos para el uso y los aplicamos prestamente a los casos individuales... sin invertir el proceso mediante el cual se consiguieron, o sin evocar o disponer ante nuestras mentes las numerosas e intrincadas consideraciones de que son *abreviaturas manuales* [cursivas mías]... La verdadera teoría es un *compendio* de verdades particulares... Hablando, pues, de manera general, la conducta humana está inevitablemente *guiada* [cursivas mías] por *reglas* o por *principios* o *máximas* (*ibid.*, pp. 117-118). No es preciso que me detenga a mostrar cómo estas observaciones se inclinan a la mira sumaria. Más adelante, cuando Austin viene a tratar de casos de 'ocurrencia comparativamente rara', sostiene que las consideraciones específicas pueden sobreponerse a las generales. 'Si observamos las razones de donde hemos inferido la regla, sería absurdo que las tuviéramos por inflexiones. Hemos de hacer a un lado la *regla*, consiguientemente, acudir por lo directo al *principio* según el cual están cortadas nuestras reglas y calcular las consecuencias *específicas*, cuanto nuestro conocimiento y capacidad lo permitan' (*ibid.*, pp. 120-121). El punto de vista de Austin es interesante porque muestra cómo se puede uno acercar a la concepción de la práctica y luego apartarse de ella.

En *A System of Logic*, libro VI, cap. xii, par. 2, Mill distingue claramente entre la posición del juez y la del legislador, y al proceder así quiere dar a entender que existe distinción entre los dos

rrir, se debería aplicar directamente el principio utilitarista, caso por caso, y de nada servirían las reglas que hablaran de decisiones pasadas.

2. Las decisiones hechas sobre los casos particulares, lógicamente son anteriores a las reglas. Como

conceptos de reglas. Sin embargo, distingue las dos posiciones, para ilustrar la diferencia existente entre los casos en que hay que aplicar una regla que gobierne la conducta subsiguiente. Es el último caso el que le interesa y toma la 'máxima del procedimiento' del legislador como típica de lo que son reglas. En el par. 3, queda bien clara la mira sumaria. Por ejemplo, dice de las reglas de conducta que se han de tomar como provisionales, puesto que están hechas para los casos que más abundan. Dice que 'señalan' la manera como es menos peligroso actuar; sirven como 'admonición' de que se ha encontrado un modo de conducta que concuerda con los casos más comunes. En *Utilitarism*, cap. ii, par. 24, aparece la concepción sumaria en la respuesta de Mill a la misma objeción que trató Austin. Aquí habla de las reglas como 'corolarios' del principio de la utilidad; estas reglas 'secundarias' se comparan a 'hitos' y 'mojones'. Se basan en larga experiencia, por lo que hacen innecesaria la aplicación del principio utilitarista a cada caso. En el par. 25, Mill se refiere al cometido del principio utilitarista consistente en adjudicar entre las reglas morales competentes. Habla aquí como si se aplicara el principio utilitarista directamente al caso particular. En la mira de la práctica, se ha de emplear el principio más bien en determinar cuál de las maneras es mejor para hacer que la práctica sea consistente. Se ha de advertir que mientras en el par. 10 la definición de Mill respecto del utilitarismo hace aplicación del principio de utilidad a la moralidad, es decir, a las reglas y preceptos de la conducta humana, la definición del par. 2 emplea la frase 'las acciones son correctas en la *proporción* en que tienden a promover la felicidad' [cursivas mías], y esto inclina hacia la mira sumaria. En el último párrafo del ensayo 'On the Definition of Political Economy', *Westminster Review* (octubre de 1336), Mill dice que sólo en el arte, en contraposición a la ciencia, se puede hablar propiamente de excepciones. En cuestiones de prácticas, si algo es lo que se suele hacer 'en la mayoría de los casos', entonces se convierte en regla. 'Al tratar de arte podemos hablar, *sin que quepa objetar*, de *regla* y de *excepción*, entendiendo por regla los casos en que existe una preponderancia ..de inducciones para actuar de una manera particular, y por excepción, de los casos en que la preponderancia está en el caso contrario'. Estas observaciones sugieren también la mira sumaria.

En *Principia Ethica* de Moore, cap. v, hay una discusión complicada y difícil de las reglas morales. No la examinaré aquí, salvo para expresar la sospecha de que prevalece la concepción sumaria. No hay duda de que Moore habla frecuentemente de la utilidad de las reglas cuando se suelen seguir, y de las acciones cuando se suelen practicar, pero es imposible que estos pasajes cuadren en

las reglas tienen su razón de ser en la necesidad de aplicar el principio utilitarista a muchos casos similares, se sigue que un caso particular (o los distintos casos que se le asemejen) puede existir independientemente de que haya una regla para ese caso. Hay, pues, casos particulares anteriores a la existencia de una regla que los abarque, pues sólo si nos encontramos con un número de casos de determinado tipo podremos formular una regla. Así, podemos describir un caso particular como uno del género requerido, independientemente de si existe una regla que ataña a *ese* género de caso. Dicho de otra manera: aquello a lo que se refieren las *Aes* y las *Bes*, en reglas de la forma 'Siempre que *A* hace *B*', se puede describir como *Aes* y *Bes*, independientemente de si existe una regla 'Siempre que *A* hace *B*', o independientemente de que exista un cuerpo de reglas que constituya una práctica de la que esa regla es una parte.

Como ilustración de lo anterior, consideremos una regla o máxima que pudiera surgir de esta manera: supongamos que una persona trata de decidir si debe revelar a alguien, irremediablemente enfermo, cuál es la enfermedad que tiene, cuando se le ha pedido que se lo diga. Supongamos que la persona, reflexionando, resuelva, por motivos utilitaristas, que no le ha de revelar la verdad; y supongamos también que, por esta y otras ocasiones, formula una regla referente a no decir la verdad cuando alguien deshauciado le pregunte qué tiene. Hay que advertir

la noción estadística de la generalidad que admite la concepción sumaria. Esta concepción viene sugerida por el hecho de que Moore toma el principio utilitarista como si aplicara directamente a las acciones particulares (pp. 147-148) y por la noción que tiene de que una regla es algo que indica cuál, de unas cuantas opciones, es la que tiene más probabilidad de aportar el mayor bien total, a cualquiera, en el futuro inmediato (p. 154). Habla de la 'ley ética' como predicción, como generalización (pp. 146, 155). La concepción sumaria es la que se pergeña en su discusión de las excepciones (pp. 162-163) y de la fuerza de los ejemplos de infracciones de reglas (pp. 163-164).

que cosas como cuando alguien sin remedio pregunte sobre su enfermedad y alguien se la revele, se pueden describir como tales, exista o no esa regla. La ejecución de la acción a que se refiere la regla no requiere el escenario de una práctica de que sea parte esa regla. Esto es lo que he querido decir al afirmar que, en la mira sumaria, los casos particulares son anteriores lógicamente a las reglas.

3. En principio, toda persona tiene derecho a reconsiderar la corrección de una regla y a preguntarse si es conveniente o no seguirla en el caso particular. Como las reglas son guías y ayudas, cabe preguntarse si en las decisiones pretéritas no se incurrió en algún error en la aplicación del principio utilitarista para formar la regla en cuestión y si es o no es lo mejor en tal caso. La razón de las reglas es que la gente no es capaz de aplicar el principio utilitarista sin mayor esfuerzo e impecablemente; es preciso ahorrar tiempo y plantar un jalón. Según esto, una sociedad de utilitaristas racionales carecería de reglas y cada individuo aplicaría el principio utilitarista directamente y sin roces, con acierto y caso por caso.

Por otra parte, nuestra sociedad formula reglas como guías para alcanzar esas decisiones idealmente racionales en casos particulares, guías que se han formado y probado al socaire de las experiencias de generaciones. Si se aplica a las reglas esta manera de ver, aparecen como máximas, como 'reglas de buen cubero', y es de dudar si hay algo a lo que se aplique la concepción sumaria y ésta pueda continuar llamándose *regla*. Discutir en filosofía como si las reglas fueran así es incurrir en un error.

4. El concepto de regla *general* toma la siguiente forma. Se supone que uno estima en qué porcentaje de casos probables se puede confiar en una regla porque exprese la resolución correcta, esto es, la decisión a que se llegaría si se aplicara el principio utilitarista correctamente y caso por caso. Si se estima que en la mayoría de casos la regla dará la decisión apropiada, o si se estima que la probabilidad

de cometer una equivocación al aplicar el principio utilitarista directamente por sí mismo es mayor que la probabilidad de cometer un error por seguir la regla —y si estas consideraciones son las que hace en general la gente—, entonces se justificaría su adopción como regla general. De esta manera se puede dar razón de las reglas *generales* en la mira sumaria. Sin embargo, también tendrá sentido hablar de la aplicación caso por caso del principio utilitarista, pues fue porque se trató de prever los resultados de hacer tal cosa como se consiguieron las apreciaciones iniciales sobre las que depende la aceptación de la regla. El que se está tomando una regla de acuerdo con la mira sumaria se verá por la naturalidad con que se hable de la regla, como guía o como máxima, o como una generalización de la experiencia, o como algo que se ha de dejar de lado en casos extraordinarios donde no hay seguridad de que la generalización cuadre, por lo que el caso se ha de tratar según sus méritos. Así, con este concepto va la noción de la excepción particular que convierte a una regla en sospechosa en una contingencia especial.

La otra concepción de las reglas la denominaré concepción de la práctica; según esta mira, las reglas vienen a definir una práctica. Las prácticas se instituyen por distintas razones, pero una de ellas es porque, en muchos sectores de la conducta, si cada persona tuviera que decidir caso por caso qué hacer según principios utilitaristas, se crearía confusión, y porque los conatos de coordinar la conducta previendo cómo actuarán los demás parecen no resultar. Como alternativa, uno se da cuenta de que lo requerido es sentar una práctica, especificar una nueva forma de actividad, y de aquí se ve que la práctica supone necesariamente abdicar la libertad plena para actuar sobre bases utilitaristas y prudenciales. Es marchamo de una práctica que el ser iniciado en ella exija saber de las reglas que la definen y que se recurra a dichas reglas para corregir el comportamiento de quienes se relacionan con ellas. Quienes

siguen una práctica aceptan las reglas como definitorias de ella. Las reglas no se pueden tomar cual si describieran simplemente cómo se comportan quienes siguen la práctica; no es que actúen sin más como si estuvieran obedeciendo las reglas. Así, es esencial en la noción de práctica que las reglas se conozcan públicamente y se conozcan como definitivas, y es esencial también que las reglas de una práctica se puedan enseñar e imponer para que rindan una práctica coherente. Según esta concepción, pues, las reglas no generalizan las decisiones de los individuos que aplican el principio utilitarista directa e independientemente a los casos particulares que se van presentando. Por el contrario, las reglas definen una práctica y en sí son sujeto del principio utilitarista.

Para mostrar las diferencias importantes entre esa manera de encuadrar las reglas en la teoría utilitarista y la manera anterior, consideraré las diferencias entre las dos concepciones según los puntos antes tratados.

1. En contraposición a la mira sumaria, las reglas de las prácticas son anteriores, lógicamente, a los casos particulares. Esto es así porque no puede darse el caso particular de una acción que caiga bajo la regla de una práctica, a menos que exista la práctica. Esto se aclarará mejor como sigue: en una práctica hay reglas que instauran oficios, especifican ciertas formas de acción apropiadas para los distintos oficios y fijan penalidades por el quebranto de las reglas, etc. Podemos pensar que las reglas de una práctica definen los oficios, las jugadas y las ofensas. Ahora, lo que se indica al decir que la práctica es anterior lógicamente a los casos particulares es lo siguiente: dada cualquier regla que especifique una forma de acción (jugada), no se *describirá* como tal tipo de acción a aquel proceder que se supone cae bajo esa regla, si concedemos que existe la práctica, a menos que efectivamente exista tal práctica. En el caso de acciones especificadas por prácticas es lógicamente imposible llevarlas a cabo fuera del escena-

rio dispuesto por esas prácticas, pues a menos que exista la práctica y a menos que se cumplan las propiedades requeridas, lo que uno haga, lo que uno juegue, no entrará como forma de acción que la práctica específica. Lo que uno haga se describirá de alguna *otra* manera.

Se puede ilustrar este punto a partir del juego del baseball. Muchas de las acciones que se realizan en el juego del baseball se pueden efectuar por sí propio o por otros, haya o no juego de baseball. Por ejemplo, se puede lanzar la pelota, correr o blandir un pedazo de madera de cierta forma; pero no es posible robarse una base, retirar al bateador, pasar a primera base, cometer un error o impedir ganar una base, aunque se pueden hacer ciertas cosas que parezcan asemejarse a esas acciones, como robar una base, perderla, etc. Retirar a un jugador, robar una base, impedir la entrada en ella, etc., son acciones que sólo pueden ocurrir en un juego. Independientemente de lo que haga una persona, sus actos no se pueden describir diciendo que entra en base, falla o entra en primera, a menos que se les puedan describir así jugando ella al baseball, y para hacer esto se exige la práctica regulada, que es lo que constituye el juego. La práctica es anterior, lógicamente, a los casos particulares: a menos que exista la práctica, carecen de sentido los términos que se refieren a acciones especificadas por ella [21].

[21] Alguien creerá que es un error decir que una práctica es anterior lógicamente a las formas de acción que especifica, basándose en que si no hubiera ejemplos de acciones que caen bajo una práctica, entonces nos sentiríamos fuertemente inclinados a decir que tampoco había práctica alguna. Los diseños de una práctica no constituyen práctica. El que haya una práctica supone que haya ejemplos de gente que la ha practicado y que la practica (con los debidos matices). Esto es correcto, pero no empece que cualquier ejemplo particular de una forma de acción especificada por una práctica presuponga la práctica. Esto no es así según la mira sumaria, puesto que cada ejemplo tiene que estar 'allí' antes que las reglas, por así decir, como algo de donde se extrae la regla aplicando directamente el principio utilitarista.

2. La mira de la práctica conduce a una concepción por completo diferente de la autoridad que cada persona tiene para decidir sobre la conveniencia de seguir una regla en casos particulares. Seguir la práctica, ejecutar aquellas acciones especificadas por ella, equivale a seguir las reglas convenientes. Si alguien quiere realizar una acción que especifica cierta práctica, entonces no hay otro medio si no es siguiendo las reglas que la definen. Por tanto, no tendrá sentido que alguien se pregunte si una regla de una práctica se aplica correctamente a *su* caso cuando la acción que está contemplando es de una forma definida por una práctica. Si alguien preguntara tal cosa, demostraría simplemente que no entendió la situación en la que estaba actuando. Si alguien desea efectuar una acción especificada por una práctica, la única pregunta legítima se refiere a la naturaleza de la práctica en sí ('¿Cómo he de hacer el testamento?').

Este particular se ilustra con la conducta que se puede esperar de un jugador en el juego. Si se desea jugar un juego, no se tratan las reglas del juego como guías sobre qué es lo mejor en casos particulares. En el baseball, si un bateador preguntara '¿Se me concederán cuatro *strikes*?', se supondría que pregunta cuál es la regla y, una vez que se le hubiera dicho cuál es ésta, si dijera que quería decir que en esa ocasión piensa que lo mejor para él es tener cuatro *strikes* en vez de tres, se tomaría como una broma. Alguien puede aducir que el baseball sería un juego mejor si se permitieran cuatro *strikes* en vez de tres, pero no es posible imaginar que las reglas sean guías respecto de lo que es mejor en total en los casos particulares, y cuestionar su aplicabilidad a los casos particulares como casos particulares.

3 y 4. Completando los cuatro puntos de comparación con la mira sumaria; es claro por lo que se ha dicho que las reglas de las prácticas no son guías que ayuden a decidir correctamente los casos particulares, cual juzgados por algún principio ético superior.

Y ni la noción cuasiestadística de generalidad ni la noción de excepción particular pueden aplicarse a las reglas de las prácticas. Será regla más o menos general de una práctica aquella que, de acuerdo con la estructura de la práctica, se aplique a más o menos clases de casos que se desprendan de ella, o deberá ser una regla más o menos básica para el entendimiento de la práctica. De nuevo, un caso particular no puede ser excepción a una regla de la práctica. La excepción es más bien una cualificación o una especificación ulterior de una regla.

Se sigue de lo que hemos dicho acerca de la concepción de la práctica que si se pregunta a una persona que ejercita una práctica por qué hace *ella lo* que hace, o si se le dice que defienda lo que hace, entonces su explicación o defensa estribará en remitir al interrogante a la práctica en cuestión. No puede decir de *su* acción, si es una acción especificada por una práctica, que lleva a cabo esa acción y no otra porque piensa que es lo mejor en total [25]. Cuando se interroga a un hombre que sigue una práctica por qué actúa así, éste ha de suponer que el preguntante o bien no sabe de qué se trata ('¿Por qué tanta prisa en pagarle?' 'Le prometí pagarle hoy') o no sabe cuál es la práctica. No se trata tanto de justificar la acción particular como de explicar o mostrar que está de acuerdo con la práctica. La razón está en que va contra la escenificación de la práctica que la acción particular de uno se describa como es. Sólo se puede *decir* qué es lo que uno está haciendo remitiéndose a la práctica. Para explicar o defender la acción propia, como acción particular, se la hace encajar en la práctica que la define. Si no se acepta esto, es señal de que se está preguntando algo distinto, referente a si alguien está justificado en aceptar o tolerar la práctica. Cuando lo que se cuestiona es la práctica,

25 Una charada filosófica (en boca de Jeremy Bentham): 'Cuando corro al otro wicket, luego que mi compañero ha dado un buen tiro, lo hago porque es lo mejor en total.'

acudir a las reglas (decir cuál es la práctica) de nada sirve; pero cuando lo que se cuestiona es la acción particular definida por la práctica, no se puede hacer nada más que remitirse a las reglas. Respecto de las acciones particulares, quien no sepa bien de qué práctica se trata o si desconoce que hay que seguirla, sólo tiene una pregunta que hacer. Esto se ha de contraponer al caso de la máxima, que puede tomarse como atinada en esta ocasión, cual si se decidiera por *otros* motivos, lo que en cierto sentido es un reto al caso, porque se cuestiona si estos otros motivos apoyan en efecto la decisión al respecto.

Si se comparan las dos concepciones de reglas que he tratado se puede ver que la concepción sumaria pasa por alto la importancia de la distinción entre justificar una práctica y justificar las acciones que caen bajo ella. Según este modo de ver, las reglas se consideran como guías cuya fidelidad es indicar la decisión idealmente racional sobre el caso particular dado, que rendiría la aplicación inmaculada del principio utilitarista. En principio se tiene la plena opción de utilizar las guías o de descartarlas, como lo avale la situación, sin que el oficio moral personal se altere en modo alguno: se descarten las reglas o no, la persona mantiene siempre el oficio de individuo racional que busca, caso por caso, realizar lo mejor en su totalidad. Pero en la concepción práctica, si alguien mantiene un oficio definido por una práctica, entonces las cuestiones referentes a las acciones propias en ese oficio se dirimen remitiéndose a las reglas que definen la práctica. Si alguien busca poner en duda esas reglas, el oficio particular sufre un cambio fundamental: entonces se presume que el propio oficio tiene poder para cambiar y criticar las reglas, o que se trata del oficio de un reformador, etc. La concepción sumaria se desentiende de la distinción de oficios y de las distintas formas de argumentación que les son propias. Según tal concepción existe un oficio y no varios oficios. Por tanto, obnubila el hecho de que el principio utilitarista, en el caso de

acciones y oficios definidos por alguna práctica, debe aplicarse a la práctica de modo que los argumentos generales utilitaristas no estén al alcance de aquellos que actúan en los oficios así definidos [26].

En lo que he dicho se necesitan algunas calificaciones. En primer lugar, puede haber parecido que he hablado de la concepción sumaria y de la práctica de las reglas como si sólo una de ambas fuera verdadera, y que si era verdadera para cualquier regla, entonces tenía que ser verdadera para *todas* las reglas. Es claro que no he querido decir tal cosa. (Son los críticos del utilitarismo quienes cometen este error, si sus argumentaciones contra el utilitarismo presuponen una concepción sumaria de las reglas de las prácticas.) Algunas reglas encajarán en una concepción y otras en otra; y así existen reglas de prácticas (reglas en sentido estricto), máximas y 'reglas de buen cubero'.

En segundo lugar, existen ulteriores distinciones válidas para clasificar las reglas, distinciones que deberían llevarse a cabo si se consideraran otras cuestiones. Las distinciones que he deslindado son las más pertinentes a asunto tan especial como el que he tratado y no llevan la intención de ser exhaustivas.

Por fin, habrá muchos casos limítrofes en los que será difícil, si no imposible, decidir cuál es la concep-

26 ¿Cómo se aplican estas observaciones al caso de la promesa sólo conocida por el padre y el hijo? Bien, a primera vista el hijo ciertamente hace las veces de prometiente, y —según es práctica— no puede sopesar el caso según bases generales utilitaristas. Supongamos, en cambio, que desee considerarse en el papel de alguien con título para criticar y cambiar la práctica, dejando de lado la cuestión respecto del derecho de pasar de su oficio previamente asumido, a otro. Entonces puede considerar los argumentos utilitaristas como aplicados a la práctica; pero en cuanto haga esto, verá que existen argumentos que no le permitirán la defensa general utilitarista en la práctica de esta clase de caso, pues proceder así imposibilitaría pedir y conceder un tipo de promesa que con frecuencia se desea estar en disposición de pedir y de conceder. Por tanto, no ha de desear cambiar la práctica y, en consecuencia, como prometiente no tiene otra opción sino cumplir la promesa.

ción de las reglas aplicables. En todo concepto existen tales casos limítrofes, pero han de abundar con conceptos como los de práctica, institución, juego, regla, etc. Wittgenstein ha mostrado cuán fluidas son esas nociones[27]. Lo que he hecho es recalcar y deslindar dos nociones correspondientes al propósito restringido de este escrito.

IV

Lo que he tratado de mostrar al distinguir entre dos concepciones de reglas es que existe un modo de considerar las reglas que permite la opción de estimar los casos particulares según bases generales utilitaristas, mientras que existe otra concepción que no admite tal posibilidad, excepto hasta el punto en que las mismas reglas lo autoricen. Quiero señalar que la tendencia en filosofía a imaginar las reglas de acuerdo con la concepción sumaria puede haber cegado a los filósofos morales la opción de ver la trascendencia de la distinción entre justificar una práctica y justificar una acción particular que cae bajo ella, y ello debido a que se trastoca la fuerza lógica de la referencia a las reglas, en el caso de que haya ataque contra una acción particular que caiga bajo una práctica, y porque se oscurece el hecho de que donde existe una práctica, es la práctica misma la que ha de ser el sujeto del principio utilitarista.

No es casualidad, sin duda alguna, que dos de los casos que son piedra de toque del utilitarismo, el castigo y las promesas, sean casos claros de prácticas. Bajo la influencia de la concepción sumaria es natural suponer que los funcionarios de un sistema penal, y quien haya hecho una promesa, pueden decidir

27 *Philosophical Investigations* (Oxford, 1953), i, pars. 67-71, por ejemplo.

qué hacer en casos particulares partiendo de bases utilitaristas. No se logra ver que es incompatible con el principio de práctica el arbitrio general para decidir sobre casos particulares según directrices utilitaristas, y que la discreción que uno pueda tener se define, asimismo, por la práctica (*v. g.*, un juez puede tener arbitrio para determinar la pena, dentro de ciertos límites). Las objeciones tradicionales contra el utilitarismo que he discutido presuponen la atribución a los jueces, y a los que han prometido, de plenitud de autoridad moral para decidir sobre bases utilitaristas respecto de los casos particulares. Pero una vez que se ensamblan el utilitarismo y la noción de práctica, y se para mientes en que el castigo y las promesas son prácticas, se ve entonces que lógicamente queda cancelada esa atribución.

Que el castigo y la promisión son prácticas está fuera de toda duda. En el caso de la promisión se muestra esto por el hecho de que la forma de las palabras 'yo prometo' es una expresión ejecutoria que presupone la escenificación de la práctica y las propiedades definidas por ella. La expresión de las palabras 'Yo prometo' constituirá promesa sólo si existe la práctica. Sería absurdo interpretar las reglas sobre la promisión de acuerdo con la concepción sumaria. Es absurdo decir, por ejemplo, que la regla sobre que se han de cumplir las promesas ha podido surgir porque se ha visto por otros casos que es mejor en conjunto cumplir las promesas hechas; pues a menos que exista de antemano el sobreentendido de que se cumplen las promesas como parte de la práctica misma, no podrían existir casos de promesas en modo alguno.

Se ha de conceder, es claro, que las reglas que definen la promisión no están codificadas y que el concepto de lo que son depende necesariamente de la educación moral personal. Por ende, es obvio que exista considerable variación sobre cómo la gente entiende la práctica, además de amplio campo para disponer la argumentación del mejor modo posible.

Por ejemplo, según sea el trasfondo de la gente habrá diferencias sobre cuán estrictamente se hayan de tomar las defensas o de cuáles de entre ellas se puede echar mano. Pero independientemente de estas variaciones, pertenece al concepto de práctica del prometer que la defensa utilitarista general no esté al alcance del prometiente. El que esto sea así da razón de la fuerza de la objeción tradicional que he tratado. Y lo que quiero dejar en claro es que cuando se yuxtaponen el punto de vista utilitarista y la concepción de la práctica de las reglas, como se debe hacer en los casos apropiados, entonces no aparece nada en tal manera de ver que implique que deba existir tal defensa, sea en la práctica del prometer o en cualquier otra práctica.

El castigo es también un caso claro. Existen muchas acciones, en la secuencia de acontecimientos que constituye el que uno sea castigado, que presuponen una práctica. Se puede ver esto examinando la definición de castigo que di al tratar de la crítica que Carritt hace sobre el utilitarismo. La definición que allí planteé se refiere a cosas como los derechos normales del ciudadano, las reglas de la ley, el proceso a seguir en la ley, en los juicios y tribunales, en los estatutos, etc., ninguno de los cuales puede existir si no está estatuido el escenario bien elaborado del sistema legal. Sucede también que muchas de las acciones por las que se castiga a la gente presuponen prácticas. Por ejemplo, se castiga el robo, la transgresión y cosas parecidas; lo que presupone la institución de la propiedad. Es imposible decir qué es el castigo, o describir un ejemplo particular de él, sin hacer referencia a los oficios, acciones, y ofensas especificadas por las prácticas. El castigo es una tirada de un juego legal complicado y presupone el complejo de prácticas que constituyen el orden legal. Lo mismo vale para ciertos castigos menos formales: los padres, profesores, o alguien con la debida autoridad, pueden castigar a un niño, pero nadie más puede hacerlo.

Existe una interpretación equivocada de lo que he estado diciendo, sobre la que vale la pena advertir. Alguien puede pensar que el empleo que hago de la distinción entre justificar una práctica y justificar las acciones particulares que caen bajo ella compromete a uno en una actitud política y social definidas, lo que lleva a una especie de conservadurismo. Puede parecer que digo que, para cada persona, las prácticas sociales de su sociedad suministran el estándar de justificación de sus acciones; por lo tanto, que cada uno se ajuste a ellas y su conducta quedará justificada.

Esta interpretación está del todo equivocada. Lo que he tratado es más bien un asunto lógico. Es claro que posee consecuencias en asuntos de teoría ética, pero en sí no conduce a ninguna actitud particular social o política. Simplemente, cuando una forma de acción está especificada por una práctica, no existe justificación posible de la acción particular de una persona determinada, salvo haciendo referencia a la práctica. Por lo tanto, en esos casos la acción es lo que es, en virtud de la práctica, y explicarla es referirse a la práctica. No se puede derivar inferencia alguna respecto de si se han de aceptar las prácticas de la propia sociedad o no. Se puede ser tan radical como se quiera, pero en el caso de acciones especificadas por las prácticas, los objetos del radicalismo propio tienen que ser las prácticas sociales y su aceptación por la gente.

He tratado de mostrar que cuando reunimos el punto de vista utilitarista y la concepción de la práctica respecto de las leyes, cuando es apropiada esta concepción [23], podemos formularla de una manera

23 Como he dicho ya, no es fácil discernir dónde propiamente encaja esa concepción. Tampoco intento discutir en este punto las clases generales de casos a que se aplica, salvo que no se ha de dar por sentado que es aplicable a muchas de las llamadas 'reglas morales'. Tengo la sensación de que relativamente son pocas las acciones de la vida moral que se definan por las prácticas y que la concepción de la práctica es más apropiada para entender argumen-

que la salva de distintas objeciones tradicionales. He tratado de mostrar, además, cómo la fuerza lógica de la distinción entre justificar una práctica y justificar una acción que cae bajo ella se relaciona con la concepción de la práctica respecto de las leyes, y no se puede entender si se considera que las reglas de las prácticas están de acuerdo con la mira sumaria. Por qué, al hacer filosofía, se pueden considerar de esa forma, es algo que no he tratado. Las razones de esto son a todas vistas muy profundas y requerirían otro artículo.

tos legales y de estilo legal, que para el género más complejo de los argumentos morales. El utilitarismo se ha de hacer encajar en las distintas concepciones de las reglas, según sea el caso, y no hay duda de que no lograrlo ha ocasionado dificultades para su interpretación correcta.

XII

UTILITARISMO EXTREMO Y RESTRINGIDO[1]

J. J. C. Smart

De *Philosophical Quarterly*, vol. 6 (1956), pp. 344-51. Reimpreso, con enmienda, con permiso del autor y de *Philosophical Quarterly*.

I

Utilitarismo es la doctrina que enseña que la bondad de las acciones se ha de juzgar por sus consecuencias. ¿Qué entendemos aquí por 'acciones'? ¿Nos referimos a las acciones particulares o a las clases de las acciones? Según sea como interpretemos la palabra 'acciones' tenemos dos teorías diferentes, las cuales ambas merecen el apelativo de 'utilitaristas'.

(1) Si por 'acciones' entendemos acciones particulares e individuales, tenemos la doctrina sostenida por Bentham, Sidgwick y Moore. Según esta doctrina, probamos las acciones individuales por sus consecuencias, y las reglas generales como 'hay que cumplir las promesas' son reglas de buen cubero que empleamos para no tener que estimar cada vez las con-

1 Basado en artículo leído ante la Rama Victoriana de la Asociación Australasiana de Psicología y Filosofía (Victorian Branch of the Australasian Association of Psychology and Philosophy), octubre de 1955. [Este artículo se discute en H. J. McCloskey, 'An Examination of Restricted Utilitarism', *Philosophical Review* (1957), y en el libro de D. Lyons *Forms and Limits of Utilitarism* (Clarendon Press, Oxford, 1965), E.]

secuencias probables de nuestras acciones. Lo correcto o errado del cumplimiento de una promesa, en una ocasión particular, dependerá sólo de la bondad o de la maldad de las consecuencias de cumplir o quebrantar la promesa en esa ocasión particular. Es claro que parte de las consecuencias de quebrantar la promesa, parte a la cual adscribiremos de ordinario importancia decisiva, será el debilitamiento de la fe en la institución de la promisión. Sin embargo, si la bondad de las consecuencias de violar la regla *in toto* es mayor que la bondad de cumplirla, entonces debemos quebrantar la regla, independientemente de si la bondad de las consecuencias de que *cada uno* obedezca la regla sea o no sea mayor que la bondad de las consecuencias de que *cada uno* la quebrante. Dicho brevemente, no importan las reglas, salvo *per accidens* como reglas de buen cubero, y *de facto* como instituciones sociales con las que el utilitarista ha de contar al estimar las consecuencias. Llamaré a est doctrina 'utilitarismo extremo'.

(2) Se ha ido aceptando últimamente otra forma más modesta de utilitarismo. Esta doctrina se encuentra en el libro de Toulmin *The Place of Reason in Ethics,* en *Ethics* de Nowell-Smith (aunque me parece que este autor tiene escrúpulos), en *Lectures on Jurisprudence* (Conferencia II) de John Austin, e incluso en J. S. Mill, si la interpretación que de él hace Urmson (*Philosophical Quarterly,* vol. 3, pp. 33-39, 1953) es atinada. Parte de su encanto está en que parece resolver la disputa de filosofía moral entre los intuicionistas y los utilitaristas de manera que es muy clara. Los filósofos arriba citados sostienen, o parecen sostener, que las reglas morales son más que reglas de buen cubero. En general, la corrección de una acción *no* se ha de juzgar valorando sus consecuencias, sino por la consideración de si cae o no bajo cierta regla. El que una regla se haya de considerar como regla moral aceptable se ha de decidir, sin embargo, considerando las consecuencias de aceptar la regla. Dicho latamente, las acciones se han de

juzgar por las reglas y las reglas por sus consecuencias. Los únicos casos en que hemos de sopesar la acción individual directamente por sus consecuencias son (*a*) cuando la acción aparece bajo dos reglas diferentes, una de las cuales la secunda y la otra la prohibe, y (*b*) cuando no hay regla alguna que gobierne el caso. A esta doctrina la denominaré 'utilitarismo restringido'.

Se ha de advertir que la distinción que hago ataja y difiere del todo de la distinción que comúnmente se hace entre el utilitarismo hedonista e ideal. Bentham fue ejemplo de utilitarista hedonista extremo, mientras que Moore lo fue de utilitarista ideal, a la vez que Toulmin (quizá) podría ser clasificado como utilitarista ideal extremo. El utilitarista hedonista sostiene que la bondad de las consecuencias de una acción es sólo función de su placibilidad, mientras que el utilitarista ideal —como Moore— defiende que la placibilidad no es ni siquiera condición necesaria de su bondad. Parece que Mill, si hemos de tomar en serio sus observaciones sobre placeres superiores e inferiores, no es ni hedonista puro ni utilitarista ideal puro. Parece propugnar que la placibilidad es condición necesaria para la bondad, pero que, además, ésta es función de otras cualidades mentales. Quizá se le debería llamar utilitarista cuasi-ideal. Cuando decimos que un estado mental es bueno, pienso que expresamos algún tipo de *preferencia racional*. Cuando decimos que es placible, juzgo que damos a entender que es deleitoso, y cuando decimos que algo es placer superior, me imagino que se entiende que se puede disfrutar más verdadera o más profundamente. No sé a ciencia cierta si 'disfrutable más profundamente' no significa ni más ni menos que 'más deleitoso, aunque no lo sea a primera vista', y por lo mismo dudo de si el utilitarismo cuasi-ideal, y posiblemente también el utilitarismo ideal, no se reduciría a utilitarismo hedonista, al examinar más de cerca la lógica de palabras como 'preferencia', 'placer', 'disfrutar', 'disfrutar profundamente', etc. Por lo de-

más, sale de los propósitos de este artículo adentrarnos en esas cuestiones. Aquí sólo me incumbe la instancia existente entre utilitarismo extremo y restringido y defenderé que las dos formas de utilitarismo pueden ser o hedonistas o no-hedonistas.

La instancia entre utilitarismo extremo y restringido se puede ilustrar mediante la observación 'pero supongamos que todos hicieran lo mismo' (Cf. A. K. Stout en artículo de *The Australasian Journal of Philosophy*, vol. 32, pp. 1-29). Stout distingue dos formas del principio de universalización: la causal y la hipotética. Cuando se dice que no se ha de hacer una acción A, porque traería malos resultados si todos (o muchos) la hicieran, puede equivaler meramente a señalar que mientras la acción A si no fuera por eso sería beneficiosa, con todo, al tomar en cuenta que hacer A llevará a otra gente a hacer A también, se echa de ver que A, en sentido lato, no es realmente beneficiosa. Si se pudiera evitar esta influencia causal (como podría suceder en el caso de la promesa hecha en una isla desierta), entonces haríamos a un lado el principio de universalización. Esta es la forma causal del principio. Quien aceptara el principio de universalización en su forma hipotética se preccuparía sólo por lo que sucedería *si* todos hicieran la acción A; no le importaría si de hecho todos pudieran hacer tal acción A. Esto es, podría decir que sería malo no votar, porque tendría malos resultados si todos tomaran esa actitud, pero no le afectarían los argumentos que supusieran que mi rechazo a votar no poseería efecto alguno sobre la inclinación de los demás a votar. Haciendo uso de la distinción de Stout, podemos decir que el utilitarista extremo aplicaría el principio de universalización en su forma causal, mientras que el utilitarista restringido lo aplicaría en la forma hipotética.

¿Cómo hemos de dirimir la cuestión entre utilitarismo extremo y restringido? Ya por anticipado quiero rechazar el enfoque a medias tintas que unas veces habla de 'investigar lo que está implícito en el

sentido común moral' y otras de 'investigar cómo suele hablar la gente sobre moralidad'. No tenemos más que leer la correspondencia que aparece en los periódicos sobre la pena capital o qué se ha de hacer con Formosa, para que nos percatemos de que el sentido común moral en parte está formado de elementos supersticiosos, o moralmente malos, o de elementos lógicamente confusos. Me dirijo a gente de buen corazón y bienintencionada, por lo que espero que si nos liberamos de la confusión lógica caerán en gran parte los elementos supersticiosos y moralmente malos. Pues incluso entre la gente de buen corazón y bienintencionada es posible hallar razones supersticiosas y moralmente malas de las creencias morales. Estas razones supersticiosas y moralmente malas se esconden tras la mampara protectora de la confusión lógica. Ante individuos que no tengan confusión lógica, pero que abiertamente sean supersticiosos o malos, no puedo hacer nada; es decir, que nuestras pro-actitudes son diferentes. Así, pues, suplico se fíen de *mi* sentido común moral y apelo al *suyo propio*, y que a la vez olvidemos lo que dice ordinariamente la gente. 'La obligación de obedecer una regla', dice Nowell-Smith (*Ethics*, p. 239), '*en la opinión de la gente* (cursivas mías) no estriba en las consecuencias beneficiosas de obedecerla en ese caso particular'. ¿Qué demuestra esto? No hay duda de que la gente anda confundida aquí y que los filósofos probablemente podrán examinar la cuestión más racionalmente.

II

Para un utilitarista extremo, las reglas morales son reglas de buen cubero. En la práctica, el utilitarista extremo de ordinario guiará su conducta apelando a las reglas ('no mientas', 'no violes las promesas',

etcétera) de la moralidad del sentido común. No se debe esto a que haya algo sacrosanto en las reglas mismas, sino a que puede decir que será lo más común que actúe de una manera utilitarista extrema si no piensa como utilitarista, pues no hay duda de que con frecuencia las acciones se tienen que realizar entre prisas. Supongamos a alguien que ve que otro se está ahogando; se echa al agua y lo salva. No ha tenido tiempo de ponerse a razonar, pero de ordinario éste será el proceder que recomendaría el utilitarista extremo si ponderara el asunto. Si, en cambio, el hombre se estuviera ahogando en un río cercano a Berchtesgaden en 1938 y tuviera el bien conocido tupé negro y el bigotillo de Adolf Hitler, el utilitarista extremo, de tener tiempo, sopesaría qué probabilidad había de que aquel hombre fuera el ruin dictador, y de ser lo suficientemente probable, según razones utilitaristas extremas, dejaría que se ahogara. Sin embargo, el que salva no tiene tiempo para esto; se fía de sus instintos, se zambulle y salva al hombre. Esta confianza en sus instintos y en las reglas morales se justifica según razones utilitaristas extremas. Además, el utilitarista extremo que supiera que el que se ahogaba era Hitler, alabaría no obstante al salvador y no lo condenaría. Porque al alabar a tal hombre secunda una disposición mental valerosa y benévola y, en general, esta disposición tiene buena utilidad (a lo mejor la próxima vez será Winston Churchill). No debemos olvidar que el utilitarista extremo puede alabar acciones que sabe que están mal. Salvar a Hitler estaría mal, pero tal hecho de salvar sería miembro de una clase de acciones que, en general, están bien, y el motivo para hacer acciones de esa clase en general es beneficioso. Al considerar cuestiones de elogio o condenación, no es la conveniencia de la acción elogiada o condenada la que se pone en juicio, sino la conveniencia de la alabanza. Puede ser conveniente alabar una acción inconveniente, e inconveniente elogiar una acción conveniente.

La falta de tiempo no es la única razón por la que un utilitarista extremo puede confiar en las reglas de la moralidad del sentido común, basado en principios utilitaristas extremos. Sabe que en casos particulares en que entran sus propios intereses, probablemente sus cálculos se sesgarán en su favor. Supongamos que no le va bien en el matrimonio y piensa divorciarse. Con toda probabilidad exagerará grandemente su propia infelicidad (y posiblemente la de su esposa) y no apreciará lo sufiente el daño que perpetrará contra sus hijos al romper el hogar. También subestimará probablemente el daño que se comete al debilitar la fe en la coyunda matrimonial. Así, probablemente llegará a la conclusión utilitarista extrema correcta, si en este caso no piensa como utilitarista extremo, sino que confía en la moralidad de sentido común.

Hay otros muchos puntos sutiles que podrían señalarse en conexión con la relación existente entre el utilitarismo extremo y la moralidad de sentido común. Todos los que he anotado, y muchos más, se encontrarán en el libro cuarto, caps. 3-5, de *Methods of Ethics*, de Sidgwick. Creo que es el mejor libro jamás escrito sobre ética y que esos capítulos son los mejores del libro. Como están casi al final de un libro tan extenso se les suele pasar por alto indebidamente. Remito al lector a la exposición clásica, en Sidgwick, de la relación entre el utilitarismo (extremo) y la moralidad de sentido común. Otro punto expuesto por Sidgwick a este propósito es si el utilitarismo (extremo), al basarse sobre principios utilitaristas (extremos), propagará entre la gente el utilitarismo (extremo). Como la mayoría no es muy filosófica y no entiende de cálculos empíricos, es probable que los individuos que la forman actúen de ordinario de una manera utilitarista extrema, si no tratan de pensar como utilitaristas extremos. Ya hemos visto cuán fácil sería aplicar erradamente el criterio utilitarista extremo en el caso del divorcio. Sidgwick parece pensar que es muy probable que el

utilitarista extremo no llegue a propagar su doctrina muy por extenso. El gran peligro para la humanidad, empero, proviene hoy del nivel de la moralidad pública, no de la privada. Existe mayor peligro para la humanidad en la bomba de hidrógeno que en el incremento de la frecuencia del divorcio, por lamentable que éste sea, y no parece que exista duda alguna respecto de que el utilitarismo extremo encaja en las relaciones internacionales. Cuando Francia se retiró de la O. N. U., porque no quería que se discutiera el caso de Marruecos, dijo que estaba en sus derechos, puesto que Marruecos y Argelia pertenecían al territorio metropolitano y nada tenían que ver con las Naciones Unidas. Se trataba de un argumento legalista, si no supersticioso. No se trataba de los llamados 'derechos' de Francia o de cualquier otro país, sino de si la causa de la humanidad quedaría mejor parada si se tratara de Marruecos en la O. N. U. (No estoy diciendo que la respuesta es 'Sí'; hay buenos motivos para suponer que de tal discusión provendrían más males que bienes.) Personalmente no dudo en decir que, fundándonos en principios utilitaristas extremos, deberíamos propagar el utilitarismo extremo lo más que pudiéramos. Aunque Sidgwick tiene razones que merecen respeto para suponer lo opuesto.

El utilitarista extremo considera, pues, las reglas morales como reglas de buen cubero y como hechos sociológicos que se han de tomar en cuenta al decidir qué hacer, como se han de tomar en cuenta hechos de otra suerte, pero que en sí las reglas morales no justifican acción alguna.

III

El utilitarista restringido considera las reglas morales como más que reglas de buen cubero para abre-

viar los cálculos de las consecuencias. En general, dice, las consecuencias no tienen importancia alguna cuando decidimos qué hemos de hacer en casos particulares. En general, sólo son importantes para decidir qué reglas implican buenas razones para actuar de determinada manera en casos particulares. Esta doctrina posiblemente da buena razón de cómo el inglés irreflexivo del siglo veinte piensa a menudo sobre la moralidad, pero sin duda es monstruosa como explicación del modo de pensar respecto de la moralidad. Supongamos que hay una regla *R* y que en el 99 por 100 de los casos los mejores resultados posibles se consiguen actuando en consecuencia con *R*. Entonces *R* es una regla de buen cubero útil. Si no disponemos de tiempo o no somos lo suficientemente imparciales para tasar las consecuencias de una acción, es en extremo buena postura decidir que lo mejor es actuar de conformidad con *R*. Pero ¿no es monstruoso suponer que si *hemos* sopesado las consecuencias y si tenemos fe perfecta en la imparcialidad de nuestros cálculos, y si *sabemos* que en este caso romper *R* tendrá mejores resultados que seguirla, hemos de obedecer la regla a pesar de todo? ¿No es erigir *R* en una especie de ídolo si la respetamos, cuando violarla nos libraría —digamos— de alguna desgracia vitanda? ¿No es una forma de latría supersticiosa a la regla (de fácil explicación psicológicamente), y no el pensamiento racional de un filósofo?

Se puede aclarar esto mejor si consideramos la comparación que hace Mill de las reglas morales con las tablas del almanaque náutico (*Utilitarism*, Everyman Edition, pp. 22-23). Esta comparación de Mill la trae Urmson como prueba de que Mill era un utilitarista restringido, pero no creo que quepa tal interpretación en modo alguno. (Aunque concuerdo con Urmson en que muchas otras cosas dichas por Mill están en armonía con el utilitarismo restringido y no tanto con el extremo. Probablemente, Mill no había pensado mucho acerca de esta distinción y pro-

pugnaba el utilitarismo, restringido o extremo, contra otras formas, en nada utilitaristas, de argumentación moral.) Dice Mill: 'Nadie dirá que el arte de navegar no se funda en la astronomía porque los navegantes no se detienen a calcular con el almanaque náutico. Como son criaturas racionales se hacen a la mar de la vida con sus mentes estructuradas respecto de las cuestiones comunes de correcto y errado, lo mismo que de otras referentes a las cuestiones harto más dificultosas de lo sensato e insensato... Sea lo que sea que aceptemos como principio fundamental de moralidad, necesitaremos principios subordinados para aplicarlo'. Párese atención en que esto, tal cual, es sólo una argumentación en favor de los principios subordinados como reglas de buen cubero. El ejemplo del almanaque náutico es engañoso porque la información que trae éste es la misma, en todos los casos, que la recabable si se hiciera un cálculo largo y laborioso partiendo de los datos astronómicos originales sobre los que se funda el almanaque. Supongamos, sin embargo, que la astronomía fuera diferente. Supongamos que el comportamiento del Sol, de la Luna y de los planetas se aproximara mucho al que manifiestan ahora, pero que en algunas raras ocasiones acaecieran pequeñas irregularidades y discontinuidades, de manera que el almanaque nos diera reglas de la forma 'en el 99 por 100 de los casos en que las observaciones son tales y tales, se puede deducir que la posición es tal y cual'. Más aún, supongamos que hubiera métodos que nos permitieran, partiendo de cálculos directos y laboriosos de los datos astronómicos originales, sin usar las burdas y manuales tablas del almanaque, sacar la posición correcta en un 100 por 100 de los casos. Los navegantes podrían emplear el almanaque porque jamás tuvieran tiempo para largos cómputos y se contentaran con un 99 por 100 de probabilidad de acierto al calcular las posiciones. ¿No sería absurdo, sin embargo, si *hicieran* los cómputos directamente y viendo que estaban en desacuerdo con el almanaque, los hicieran

a un lado y se aferraran al cálculo de éste? Otro sería el caso, es obvio, si hubiera frecuentísima probabilidad de cometer errores en los cómputos directos; entonces podríamos atenernos al resultado del almanaque, aunque supiéramos que era falible, simplemente por la operación directa podría ser errónea por razón diferente, la fiabilidad del computador. Esto sería análogo al caso del utilitarista extremo que se atiene a la regla convencional, contra los dictados de sus cálculos utilitaristas, sólo porque piensa que sus cálculos probablemente adolecen de sesgos personales. Pero si el navegante estuviera seguro de sus cómputos directos, ¿no sería insensato si persistiera en seguir el almanaque? Concluyo, pues, que si alteramos nuestras suposiciones respecto de la astronomía y el almanaque (en los que no caben excepciones) y traemos el caso a la moralidad (en cuyas reglas hay excepciones), el ejemplo de Mill pierde sus visos de apoyo a la forma restringida de utilitarismo. Permítaseme decir una vez más que no estudio aquí cómo piensan de ordinario las personas sobre la moralidad, sino cómo deberían pensar. Podríamos imaginar muy bien una raza de navegantes que poseyeran una reverencia supersticiosa por su almanaque, aunque sólo estuviera acertado en un 99 por 100 de los casos y que, airados, echaran por la borda a todo aquél que mencionara siquiera el cómputo directo; pero, ¿sería racional tal comportamiento?

Consideremos ahora un tipo de caso mucho más discutido, en el cual el utilitarista extremo puede ir contra la regla moral convencional. He prometido a un amigo moribundo en una isla desierta, de la que soy rescatado luego, que miraré por que su fortuna (sobre la que tengo control) se entregue a un jockey club. Sin embargo, luego de rescatado, pienso que sería mejor dar ese dinero a un hospital, que puede hacer más bien con él. Se puede alegar que hago mal en entregar la fortuna a un hospital. Pero, ¿por qué? (*a*) El hospital puede hacer más bien con el dinero

que un jockey club; (*b*) El presente caso se diferencia de los demás casos ordinarios de promisión porque nadie más sabe de la promesa. Si quebranto la promesa lo hago en completo secreto y no coopero en nada para debilitar la fe general en las promesas. Factor, éste, que disuadiría normalmente al utilitarista extremo de violar la promesa, incluso en casos por lo demás no bonancibles, pero que aquí no han lugar; (*c*) Existirá, no hay duda, un nimio debilitamiento en mi carácter como cumplidor habitual de lo prometido y, además, pueden hacer aparición tensiones psicológicas cada vez que se me pregunte qué quiso mi amigo que yo hiciera. Pues claramente habré de decir que me hizo prometer que diera el dinero a un hospital y, como soy veraz de ordinario, esto me vendrá muy a contrapelo. Estoy muy seguro de que si me ocurriera el caso cumpliría la promesa, pero no estamos discutiendo sobre qué me harían realizar mis hábitos morales; estamos tratando de lo que debería hacer. Además, no hemos de olvidar que si incluso fuera muy racional dar el dinero al hospital, sería también muy racional que usted me castigara o condenara si llegara a descubrir la verdad (cosa muy improbable) (*v. g.*, por haber encontrado una nota en una botella llegada a la playa). Además, concedería que si fuera muy racional dar el dinero al hospital, lo sería también que usted me condenara por ello. Regresamos otra vez a la distinción de Sidgwick entre la utilidad de una acción y la utilidad de la alabanza de la misma.

A. K. Stout trata de muchas instancias como éstas en el artículo a que he hecho referencia. No es mi intención volver sobre lo mismo, especialmente porque me parece que las argumentaciones de A. K. Stout apoyan mi propio punto de vista. Será útil, sin embargo, considerar otro ejemplo que trae. Supongamos que en tiempo de mucho calor se expide un edicto prohibiendo el uso del agua para regar jardines. Yo tengo un jardín y razono que es muy seguro que la mayoría de la gente obedecerá la orden y que como

la cantidad de agua que voy a usar es exigua en sí, no causaré ningún daño si empleo el agua secretamente. Riego y se abren unas flores que alegran a diferente gente. Se puede decir que, aunque la acción quizá fue benéfica, estuvo mal y no fue limpia.

Hay distintas cuestiones que considerar. Sin duda, mi acción merece condena. Regresamos una vez más a la distinción de Sidgwick. Una acción correcta puede ser condenada racionalmente. Además de que este tipo de incumplimiento suele descubrirse. Si tengo un hermoso jardín cuando los ajenos están secos y marchitos, sólo existe una explicación. Así, si riego mi jardín estoy debilitando mi respeto por la ley y el orden, y como esto conduce a malos resultados, un utilitarista extremo aceptaría que hice mal en regar el jardín. Supongamos ahora que el caso es distinto, y que puedo guardar el asunto en secreto: existe una parte cerrada del jardín donde cultivo flores que luego despacho secretamente para una casa de mujeres ancianas. ¿Están tan seguros aún de que hice mal en regar el jardín? Sin embargo, se trata de un caso menos trascendente que el del hospital y el del jockey club. Habrá tensiones dentro de mí. El conocimiento secreto de que he quebrantado la regla me dificultará que exhorte a los demás a guardarla. Estos malos efectos psicológicos pueden no ser insignificantes: directa e indirectamente pueden conducir a un perjuicio que al menos es del mismo orden que la felicidad que las ancianas recibirán con las flores. Vese, pues, que en el punto utilitarista extremo la cuestión presenta dos flancos.

Hasta aquí he estado considerando el deber de un utilitarista extremo en una sociedad predominantemente no utilitarista. Cambia el caso si consideramos el caso del utilitarista extremo que vive en una sociedad en que cada miembro, o la mayoría de ellos, razone como él. ¿Podría regar ahora las flores? (Concedo que en el caso ponderado hubiera estado perfectamente que las regara, lo que es dudoso.) Como primera consideración, la respuesta es que no debería

hacerlo, puesto que como se trata de una situación completamente simétrica, lo que es racional para él lo es para los demás. Por tanto, mediante un argumento de *reductio ad absurdum* se vería que regar el jardín no sería racional para nadie. No obstante, un análisis más refinado muestra que el argumento anterior no está bien, aunque lo sea suficientemente para propósitos prácticos. El argumento considera a cada persona enfrentada con la elección o de regar el jardín o de no regarlo. No obstante, existe otra posibilidad, que es que a cada persona se diera cierta probabilidad de regar el jardín mediante algún procedimiento aleatorio, como echando los dados. Esto equivaldría a adoptar lo que en la teoría de los juegos se llama 'estrategia mixta'. Si pudiéramos dar valores numéricos al beneficio privado del riego del jardín y al perjuicio público causado por 1, 2, 3, etc., personas que aplicaran el riego, podríamos extraer un valor de probabilidad de regar el jardín que cada utilitarista extremo podría darse a sí mismo. Supongamos que a es el valor que cada utilitarista extrae de regar el jardín, siendo $f(1)$, $f(2)$, $f(3)$..., el perjuicio público causado por 1, 2, 3..., personas ni más ni menos que respectivamente regaran el jardín. Supongamos que p es la probabilidad que cada persona se da de regar su jardín. Entonces podemos calcular fácilmente, como funciones de p, las probabilidades de que sean exactamente 1, 2, 3, etc., personas las que rieguen su jardín. Supongamos que esas probabilidades son p_1, p_2..., p_n. Entonces todo el beneficio neto probable se puede expresar como

$$V=p_1\ [a-f(1)]+p_2\ [2a-f(2)]+\ldots p_n\ [na-f(n)]$$

Si sabemos la función $f(x)$, podemos calcular el valor de p, para el que $(dV/dp)=0$. Tal vez sería el valor de p que racionalmente podría adoptar cada utilitarista extremo. El presente argumento no estriba en una suposición, injustificada quizá, de que los valores en cuestión han de ser mensurable, pues en un caso

práctico, como el del riego del jardín, podemos suponer sin más que *p* será tan pequeña que casi llegue a cero. Sin embargo, esta argumentación es de interés para el apuntalamiento teórico del utilitarismo extremo, puesto que los críticos del utilitarismo de ordinario hacen a un lado la estrategia mixta, suponiendo equivocadamente que las únicas opciones pertinentes y simétricas son las que poseen la forma 'todos hacen X' y 'nadie hace X' [2].

Paso ahora a un tipo de caso que puede considerarse como uno de los triunfos del utilitarismo restringido. Veamos el caso de las reglas de tránsito. Se puede decir que lo que importa es que todos procedan igual, por lo que es indiferente que la regla diga 'hay que ir por la derecha' o 'hay que ir por la izquierda'. De hecho, la única *razón* existente para que en las naciones británicas se vaya por la izquierda es que tal es la regla. Aquí la regla parece ser una razón en sí para proceder de determinada manera. Quiero impugnar esto. La regla en sí no es ninguna razón de nuestras acciones. Estaría perfectamente justificado que se fuera por la derecha si: (*a*) la regla estatuyera seguir por la izquierda, y (*b*) viviéramos en un país de superanarquistas que siempre hicieran por principio lo contrario de lo que se les ordenara. Esto nos muestra que la regla no nos da razón alguna para actuar, sino que más bien es una indicación de las acciones probables de otros, lo que nos auxilia en averiguar cuál habría de ser el procedimiento más racional. Si vivimos en un país poblado no por anarquistas, sino por utilitaristas extremos no-anarquistas, esperamos que, siendo iguales las demás cosas, observarán las reglas que se les impongan. El conocimiento de las reglas nos permite predecir cuál ha de ser su conducta y armonizar nuestras acciones con las de ellos. La regla 'seguir por la izquierda' no es, pues, una *razón lógica* para actuar, sino un *dato* antropológico para planear las acciones.

2 [Este párrafo ha sido sustancialmente modificado por el autor. Ed.]

Concluyo que, en todo caso, si existe una regla *R* cuyo cumplimiento por lo general es beneficioso, pero tal que en una clase especial de circunstancias deberíamos quebrantar, entonces en esas circunstancias deberíamos quebrantar *R*. Es claro que debemos considerar todos los efectos menos obvios de quebrantar *R*, como la reducción de la fe de la gente en el orden moral, antes de llegar a la conclusión de que quebrantar *R* está correcto; de hecho, raramente podemos llegar a tal conclusión. Según el punto de vista utilitarista extremo, las reglas morales son sólo de buen cubero, pero no son malas reglas de buen cubero. Mas si *llegamos* a la conclusión de que hemos de violar la regla y si hemos sopesado nuestra propia falibilidad y exposición al sesgo personal, ¿qué otra razón nos queda para seguir la regla? Puedo entender el 'es beneficioso' como razón para actuar, pero ¿por qué lo habría de ser 'es miembro de una clase de acciones que de ordinario son beneficiosas' o 'es miembro de una clase de acciones que, como clase, son más beneficiosas que cualquier otra clase general'? Equivaldría a decir que alguien debería jugar por Australia porque todos sus hermanos han jugado por ella, o porque el equipo australiano se ha de componer por entero de la familia Harvey, ya que sería mejor que componerlo enteramente por miembros de cualquier otra familia. El utilitarista extremo no apela a sentimientos artificiales, sino sólo a sentimientos de benevolencia y ¿a qué mejores sentimientos se puede ocurrir? Es de admitir que podemos tener una actitud en pro de algo, incluso de las reglas, pero tales pro-actitudes engendradas artificialmente saben a superstición. Vayamos a la realidad, a la felicidad y a la miseria humanas, y convirtámoslas en objetos de nuestras pro-actitudes y anti-actitudes.

El utilitarista restringido puede decir que sólo habla de *moralidad*, no de cosas tales como reglas de tránsito. No sé hasta qué punto esta objeción, de ser válida, afectaría mi argumentación, pero en todo

caso respondería que, en cuanto filósofo, concibo la ética como el estudio de cuál sería la manera *más racional* de actuar. Si mi impugnante quiere restringir la palabra 'moralidad' a un sentido más estrecho, puede hacerlo. La cuestión fundamental es la racionalidad de una acción *en general*. De manera semejante, si el utilitarista restringido quisiera apelar al uso ordinario y dijera 'lo más racional sería permitir que Hitler se ahogara, pero sin duda no sería *malo* salvarlo', yo dejaría otra vez que empleara a su gusto las palabras 'bueno' y 'malo', y me atendría a 'racional' e 'irracional'. Hemos visto que sería racional alabar al salvador de Hitler, aunque hubiera sido mayormente racional dejar que éste se ahogara. En el lenguaje ordinario, empero, 'correcto' y 'equivocado' no tienen el único significado de 'lo más racional' y 'no lo más racional', sino también el de 'loable' y 'no loable'. De ordinario, a la utilidad de una acción le corresponde utilidad de su elogio, pero —como vimos— no siempre es así. Ahora bien, el lenguaje moral se esclarecería reservando por ejemplo 'correcto' para 'lo más racional' y 'bueno' como un epíteto de elogio por el motivo de donde surgió la acción. Sería más propio de un filósofo que éste aplanara las expresiones ilógicas del lenguaje moral y tratara de reformarlas, que convertirlas en tribunal de apelación para perpetuar las confusiones.

La siguiente puede ser una defensa última del utilitarismo restringido. 'Actuar benéficamente' puede considerarse en sí como una de las reglas de nuestro sistema (aunque sería raro decir que la regla estaba justificada por beneficiosa). Según Toulmin (*The Place of Reason in Ethics*, pp. 146-8), si —pongamos por caso— el 'cumplir las promesas' entra en conflicto con otra regla, podemos dirimir el caso según sus méritos, cual si fuéramos utilitaristas extremos. Si 'actúa beneficiosamente' es en sí una de nuestras reglas, entonces siempre habrá conflicto de reglas cuando seguir la regla no sea en sí beneficioso. Si esto es así, el utilitarismo restringido cae en el utili-

tarismo extremo. Mas nadie podría leer el libro de Toulmin, o el artículo de Urmson sobre Mill, sin razonar que Toulmin y Urmson son de la opinión de que han pensado en una doctrina que *no* cae en el utilitarismo extremo, sino que, por el contrario, es su perfecionamiento.

NOTAS SOBRE LOS COAUTORES

C. L. Stevenson es profesor de Filosofía en la Universidad de Michigan. Su libro más influyente, *Ethics and Language*, se publicó en 1945, pero ha escrito muchos artículos, la mayoría de ellos sobre ética, en revistas de filosofía.

G. E. Moore, muerto en 1958, fue profesor de Filosofía en Cambridge desde 1925 a 1939, y durante la guerra mundial enseñó en muchas universidades americanas. Sus escritos, que se cuentan entre los más influyentes del presente siglo, son, entre otros, *Principia Ethica* (1903), *Some Main Problems of Philosophy* (1953), *Philosophical Papers* (1959) y *Philosophical Studies* (2.ª ed., 1960).

W. F. Frankena es profesor de Filosofía en la Universidad de Michigan. Su libro *Ethics* se publicó en 1963.

P. T. Geach, que ha enseñado durante algunos años en la Universidad de Birmingham, ahora es profesor de Filosofía en la Universidad de Leeds. Están entre sus publicaciones *Mental Acts* (1960) y *Reference and Generality* (1962).

R. M. Hare, ahora profesor de Filosofía Moral de White, Oxford, antes fue miembro de Balliol College. Sus libros, *The Language of Morals* (1952) y *Fredom and Reason* (1963), han ejercido importante influencia en las elaboraciones recientes de teoría ética.

Phiilppa Foot, miembro y 'tutora' de Filosofía de Somerville College, Oxford, ha tenido a su cargo la compilación de este volumen.

John R. Searle es profesor de Filosofía en la Universidad de California, en Berkeley. Ha escrito muchos artículos en revistas filosóficas, y en la presente serie tiene a su cargo *The Philosophy of Language*.

J. O. Urmson es miembro del Corpus Christi College, Oxford, antes profesor de Filosofía en Dundee. Su libro, *Philosophical Analysis* se publicó en 1958, y tuvo a su cargo la publicación de

las últimas conferencias de J. L. Austin sobre William James, *How to Do Things with Words* (1962).

J. D. Mabbott ha sido presidente de St. John's College, Oxford, desde 1953, del que antes fue miembro. Entre sus publicaciones están *The State and the Citizen* (2.ª ed., 1952) e *Introduction to Ethics* (1966).

John Rawls, hasta hace poco profesor de Filosofía en el Massachusetts Institute of Technology, está ahora en Harvard. Han influido vastamente sus escritos sobre teoría ética.

J. J. C. Smart es Profesor 'Hughes' de Filosofía en la Universidad de Adelaide. Su libro *Philosophy and Scientific Realism* se publicó en 1963.

BIBLIOGRAFIA

(Sin contar las citas de este volumen)

I. LIBROS

Austin, J. L., *How to do Things with Words* (Clarendon Press, Oxford, 1962).

Ayer, A. J., *Language, Truth and Logic* (Gollancz, Londres, 1936; 2.ª ed., 1946).

Baier, K., *The Moral Point of View: A Rational Basis of Ethics* (Cornell University Press, Ithaca, Nueva York, 1958).

Hare, R. M., *The Language of Morals* (Clarendon Press, Oxford, 1952).

Hare, R. M., *Freedom and Reason* (Clarendon Press, Oxford, 1963).

Lyons, D., *Forms and Limits of Utilitarism* (Clarendon Press, Oxford, 1965).

Moore, G. E., *Principia Ethica* (Cambridge University Press, Cambridge, 1903).

— *Ethics* (Home University Library, Williams and Norgate, Londres, 1912).

Nowell-Smith, P. H., *Ethics* (Penguin Books, 1954. Blackwells, Oxford, 1957).

Ogden, G. K. y Richards, I. A., *The Meaning of Meaning*, Kegan Paul, Londres, 1923).

Prichard, H. A., *Moral Obligation: Essays and Lectures* (Clarendon Press, Oxford, 1949).

Ross, W. D., *The Right and the Good* (Clarendon Press, Oxford, 1930).

— *Foundations of Ethics* (Clarendon Press, Oxford, 1939).

Schilpp, P. A. (comp.), *The Philosophy of G. E., Moore* (Northwestern University Press, Evanston, 1942; 2.ª ed., 1952, Tudor Publishing Company, Nueva York).

Schlick, M., *The Problems of Ethics* (Prentice Hall, Nueva York, 1938).

Singer, M. G., *Generalisation in Ethics* (Eyre and Spottiswoode, Londres, 1963).

Stevenson, C. L., *Ethics and Language* (Yale University Press, Nueva Haven, 1945).

— *Facts and Values: Studies in Ethical Analysis* (Yale University Press, Nueva Haven, 1963).

Warnock, M., *Ethics since 1900* (Clarendon Press, Oxford, 1960).
Williams, B. y Montefiore, A., *British Analytic Philosophy* (Routledge and Kegan Paul, Londres, 1966).
Ziff, P., *Semantic Analysis* (Cornell University Press, Ithaca, Nueva York, 1960).

II. ARTICULOS

1) *Referentes a los números del I al VIII de este volumen.*

Anscombe, G. E. M. A., 'On Brute Facts', *Analysis* (1958).
Baier, K., y Toulmin, S. E., 'On Describing', *Mind* (1952).
Barnes, W., 'Ethics Without Propositions', *Proceedings of the Aristotelian Society* (1948-9).
Black, M., 'Some Questions about Emotive Meaning', *Philosophical Review* (1964).
— 'The Gap Between «Is» and «Should»', *Philosophical Review* (1965).
Diggs, B. J., 'A Technical Ought', *Mind* (1960).
Duncan-Jones, 'Good Things and Good Thieves', *Analysis* (1966).
Flew, A., 'On not deriving «ought» from «is»', *Analysis* (1964).
Findlay, J. N., 'Morality by Convention', *Mind* (1944).
Foot, P. R., 'Moral Arguments', *Mind* (1958).
— 'Goodness and Choice', *Aristotelian Society Supplementary Volume*, XXXV (1961).
Gardiner, P. L., 'On Assenting to a Moral Principle', *Proceedings of the Aristotelian Society* (1954-5).
Gewirth, A., 'Meanings and Criteria in Ethics', *Philosophy* (1963).
Hare, R. M., 'Universalisability', *Proceedings of the Aristotelian Society* (1954-5).
— 'Descriptivism', *Annual Philosophical Lecture, Henrietta Hertz, British Academy* (1963).
MacIntyre, A., 'Hume on «Is» and «Ought»', *Philosophical Review* (1959).
Montefiore, A., 'Goodness and Choice', *Aristotelian Society Supplementary Volume*, XXXV (1961).
Moore, G. E., 'Is Goodness a Quality?', *Aristotelian Society Supplementary Volume*, XI (1932). Reimpreso en *Philosophical Papers* de Moore (George Allen and Unwin, Londres, 1959).
Patton, T. E., y Ziff, P., 'On Vendler's Grammar of «Good»', *Philosophical Review* (1964).
Phillips, D. Z., 'Does it Pay to be Good?' *Proceedings of the Aristotelian Society* (1964-5).
— 'On Morality's Having a Point', *Philosophy* (1965).
Searle, J., 'Meaning and Speech Acts', *Philosophical Review* (1962).

Stevenson, C. L., 'The Emotive Meaning of Ethical Terms', *Mind* (1937). Reimpreso en *Facts and Values* de Stevenson.
— 'Persuasive Definitions', *Mind* (1938). Reimpreso en *Facts and Values* de Stevenson.
Strawson, P. F., 'Ethical Intuitionism', *Philosophy* (1949).
— 'Social Morality and Individual Ideal', *Philosophy* (1961).
Tanner, M., 'Examples in Moral Philosophy', *Proceedings of the Aristotelian Society* (1964-5).
Thomson, J. y J., 'How not to Derive «Ought» from «Is»', *Philosophical Review* (1964).
Urmson, J. O., 'On Grading', *Mind* (1950).
Vendler, Z., 'The Grammar of Goodness', *Philosophical Review* (1963).
Winch, P., 'Can a Good man be Harmed?', *Proceedings of the Aristotelian Society* (1965-6).
Wittgenstein, L., 'Lectures on Ethics', *Philosophical Review* (1965).

2) *Referentes a los números del IX al XII de este volumen.*

Anscombe, G. E. M., 'Modern Moral Philosophy', *Philosophy* (1958).
Harrod, R., 'Utilitarism Revised', *Mind* (1936).
Harrison, J., 'Utilitarism, Universalisability, and Our Duty to be Just', *Proceedings of the Aristotelian Society* (1952-3).
McCloskey, H. J., 'An Examination of Restricted Utilitarism', *Philosophical Review* (1957).
Stout, A. K., '«But Suppose Everyone did the Same»', *Australasian Journal of Philosophy* (1954).

INDICE DE NOMBRES

Abraham, L.: 94, 97n.
Aiken, H. D.: 211n.
Anscombe, G. E. M.: 25n., 147n., 160n.
Aristóteles: 102, 109-112, 114n., 124n.
Austin, J. L.: 25, 26 n., 26, 114n., 116n., 169n.
Ayer, A. J.: 13, 54n.

Barnes, W. H. F.: 54n.
Bentham: 27, 216n., 217n., 219n., 240n., 248, 250.
Black, M.: 171n.
Bradley, F. H.: 216n.
Broad, C. D.: 54n., 56, 77, 78, 84n., 85, 89.
Butler, Bishop: 90, 91.

Carnap, R.: 54n., 91.
Carritt, E. F.: 191n., 219, 220, 223, 226, 245.
Clarke, M. E.: 81.
Cross, R. C.: 116n.

Diggs, B. J.: 26n.
Duncan Jones, A.: 54n., 99n.

Falk, W. D.: 117.
Flew, A.: 151n., 171n.
Foot, Philippa, 30, 99.
Frankena, W. K.: 15, 30.
Frege, G.: 103, 114n.

Geach, P. T.: 18, 19, 26, 30, V passim.
Goodman, N.: 36n., 39n.

Hare, R. M.: 16-19, 30, 128.
Hart, H. L. A.: 154n.
Hobbes, 216n., 219 n.
Hume: 1, 12, 17, 84, 147, 148n., 151, 210n., 219n.

Jackson, R.: 188.
Joseph, H. W. B.: 114n.
Jury, G. S.: 81.

Kant, 108, 166, 183n., 189, 197.

Laird, J.: 83n., 84.
Leibniz: 189.
Lyons, D.: 210n., 248n.

Mabbott, J. D.: 27, 28, 30, 211n., 215n., 226n.
McCloskey, H. J.: 188n., 210n., 248n.
MacIntyre, A. C.: 171n.
McTaggart, J. E. M.: 94.
Melden, A. I.: 225n.
Mill, J. S.: 27, 28, 83-85, IX passim, 200, 202, 203, 204, 205, 210n., 219n., 232n., 233n., 249, 256, 265.
Moore, G. E.: 10-16, 27, 30, I passim, 81, 84-91, 102n., 190, 200, 204, 223n., 230n., 233n., 248, 250.

Nietzsche: 146.
Nowell Smith, P. H.: 211n., 226n., 249, 252.

Ogden, C. K.: 13.
Osborne, H.: 95n., 97n.

Patton, T. E.: 99n.
Perry, R. B.: 89, 90, 94, 96n.
Phillips, D. Z.: 126n.
Pickard-Cambridge, W. A.: 223n.
Platón: 124n., 146, 189.
Prichard, H. A.: 11, 102n.

Quinton, A. M.: 211n., 212n.

Radzinowicz, L.: 216n.
Ramsey, F. P.: 46.
Rawls, J.: 28-29, 166n., 208, 210n.
Richards, I. A.: 13, 50.
Robbins, L.: 230n.
Ross, Sir David: 11, 102n., 109, 111, 117, 203, 206, 216n., 222n., 223n., 224-5.
Russell, Bertrand: 54n.

Searle, J. R.: 23-25, 30, VIII passim.
Sidgwick, H.: 83, 219n., 227n., 248, 254-255, 260.
Smart, J. J. C.: 27, 28, 30.
Sorley, W. R.: 83.
Spencer, Herbert: 83, 85.
Stevenson, C. L.: 12-17, 30, 31, II passim, 117, 128.
Stout, A. K.: 251, 259.

Tanner, M.: 126n.
Taylor, A. E.: 81.
Thomson, J. F.: 151n.
Thomson, J. J.: 151n.
Tomás de Aquino: 110, 144n.
Toulmin, S. E.: 211n., 250, 264, 265.
Urmson, J. O.: 26, 28, 30, 200-5, 210n., 249, 265.
Vendler, Z.: 99n.
Westermarck, E. A.: 32n.
Wheelwright, P. E.: 85, 85n., 86n.
Whittaker, T.: 85n.
Williams, D. C.: 82.
Wisdom, J.: 93n.
Wittgenstein, L.: 107, 130n., 243.
Wood, L.: 85, 87.
Ziff, P.: 99n.

ÍNDICE GENERAL

		Págs.
Introducción		9
I.	*Argumentos de Moore contra ciertas formas de naturalismo ético*, por C. L. Stevenson	31
II.	*Réplica a mis críticos*, por G. E. Moore.	56
III.	*La falacia naturalista*, por W. K. Frankena	80
IV.	*Bien y mal*, por P. T. Geach	99
V.	*Geach: bien y mal*, por R. M. Hare ...	113
VI.	*Creencias morales*, por Philippa Foot.	126
VII.	*Cómo derivar 'debe' de 'es'*, por John R. Searle	151
VIII.	*El juego del prometer*, por R. H. Hare.	171
IX.	*La interpretación de la filosofía moral de J. S. Mill*, por J. O. Urmson	188
X.	*Interpretaciones del 'utilitarismo' de Mill*, por J. D. Mabbott	200
XI.	*Dos conceptos de reglas*, por John Rawls	210
XII.	*Utilitarismo extremo y restringido*, por J. J. C. Smart	248
Notas sobre los coautores		267
Bibliografía		269
Indice de nombres		272